New model interaction potential for the description of charged particle motion in matter

E. G. Sheĭkin

Scientific-Research Enterprise for Hypersonic Systems, 196066 St. Petersburg, Russia

A new model interaction potential, in the form of a screened Coulomb potential, is proposed. Analytical expressions are derived for the stopping power of ions in elastic collisions. A program is written for Monte Carlo calculations of the ion ranges in an amorphous substance, taking inelastic losses into account in the continuous-slowing approximation and taking elastic collisions into account in the approximation of the new model interaction potential. The ranges of Cu and Rb ions in C and B targets are calculated. The results of the calculations are in good agreement with experiment. © *1999 American Institute of Physics.* [S1063-7842(99)00105-1]

INTRODUCTION

Various models of the interatomic interaction potential are used in describing the motion of fast charged particles in matter. Extensive use has been made of screened Coulomb potentials,[1] which are written in the form

$$V(r) = \frac{Z_1 Z_2 e^2}{r} \Phi\left(\frac{r}{a}\right), \tag{1}$$

where Z_1 and Z_2 are the nuclear charges of the colliding atomic particles, e is the electron charge, r is the distance between the colliding atoms, a is the screening length, and $\Phi(r/a)$ is the screening function.

Usually $\Phi(r/a)$ is approximated by its expansion in a sum of exponentials:[1]

$$\Phi(r/a) = \sum_{i=1}^{n} c_i \exp(-d_i r/a). \tag{1a}$$

The values of the coefficients of the expansion c_i and d_i for various potentials are given in Refs. 1 and 2. The simplest potential of the type (1) is the Bohr potential, for which in Eq. (2a) we have $n=1$, $c_1=1$, and $d_1=1$. Interaction potentials of the type (1) are used in the Monte Carlo computer simulation of the sputtering of materials, the reflection of ions from the surface of a solid, ion implantation, etc. Despite the relatively simple form of the interaction potential (1) and (1a), the calculation of the parameters of a fast particle after elastic scattering by the target atoms requires numerical integration,[1] which drastically increases the computing times and involves a large number of trials. It would therefore be helpful to construct a new model potential that would correctly reflect the principal scattering laws and that could be used to derive analytical expressions for calculating the parameters of a fast particle after elastic scattering.

CONSTRUCTION OF A NEW MODEL INTERACTION POTENTIAL

Grande *et al.*[3] have published the results of a comparison of the experimental ranges of low-energy and medium-energy heavy ions in boron and carbon with the results of Monte Carlo calculations using the TRIM code. The results of calculations by the TRIM code, which uses a so-called universal interaction potential to describe the elastic scattering of ions, disagree strongly with the experimental results. On the other hand, the passage of low-energy and medium-energy heavy ions through matter has been investigated in a series of papers[4–7] using a simpler approximation of the interaction potential. In those studies a modified hard-sphere model was used, where the scattering of ions was treated in the hard-sphere model and the total elastic scattering cross section was assumed to depend on the energy of the ion; it was obtained in terms of the known energy dependences of the ion stopping power. The analytical expressions obtained in Refs. 5 and 6 are in good agreement with experimental results[3] for low ion energies corresponding to a normalized energy $\varepsilon < 0.1$ (standard notation is used here). When the energy is increased in the theory, the standard deviations of the projected ion ranges are observed to be too high, presumably because the interaction potential used in the theory is too rigid.

In light of the unquestionable success of the modified hard-sphere model in describing the passage of low-energy heavy ions through matter, it is proposed that the screening function be introduced in the simplest form, which has an effect similar to the hard-sphere model in that it ensures a finite radius of interaction of the colliding particles, but with a softer interaction potential than for hard spheres, which goes over to a Coulomb potential at high ion energies. The proposed screening function is

$$\Phi(r) = \begin{cases} 1 - r/a & \text{for} \quad r \leq a, \\ 0 & \text{for} \quad r > a, \end{cases} \tag{2}$$

where a is the screening length.

The scattering angle of the incident ion in an elastic collision with a target atom in center-of-mass coordinates (CM frame) is given, according to Ref. 1, by the relation

$$\chi = \pi - 2\varphi_0, \tag{3}$$

where

$$\varphi_0 = \int_{r_{min}}^{\infty} \frac{\rho\, dr}{r^2 \sqrt{1 - \dfrac{\rho^2}{r^2} - \dfrac{V(r)}{E_r}}},$$

ρ is the impact parameter, r_{min} is the distance of closest approach of the colliding particles, E_r is the kinetic energy in the CM frame, which is related to the ion energy E in the laboratory frame (for stationary target atoms) by the equation $E_r = m_2 E/(m_1 + m_2)$, m_1 is the mass of the ion, and m_2 is the mass of the target atom.

Substituting the potential (3) into Eq. (1) with the screening function in the form (2), after simple mathematical transformations we obtain an equation for the cosine of the scattering angle in the CM frame:

$$\cos\chi = 1 - \frac{2(1-t)(\tilde{\rho}/a)^2}{4t + 4t(\tilde{\rho}/a) + (\tilde{\rho}/a)^2}, \tag{4}$$

where $t = (\rho/a)^2$ and $\tilde{\rho} = Z_1 Z_2 e^2/E_r$.

The resulting equation (5) can be used to determine the ion stopping power S_n in elastic stopping and the straggling Ω^2 of the energy losses in elastic collisions. According to Ref. 8, these functions are given by the relations

$$S_n = \int_0^{T_{max}} T\, d\sigma, \quad \Omega^2 = \int_0^{T_{max}} T^2 d\sigma,$$

where T is the energy transferred to the target atom in collision with the ion, $T_{max} = 4 m_1 m_2 E/(m_1 + m_2)^2$ is the maximum energy transferred, and $d\sigma$ is the differential scattering cross section.

Taking into account the relations $T = T_{max}\sin^2(\chi/2)$ and $d\sigma = 2\pi\rho d\rho$, we obtain

$$S_n = T_{max}\pi a^2 f(X), \quad \Omega^2 = T_{max}^2 \pi a^2 F(X),$$

where

$$f(X) = X[(1+X)\ln(1 + 1/X) - 1],$$

$$F(X) = X^2[1 + (1+X)(1/X - 2\ln(1 + 1/X))],$$

$$X = \frac{(\tilde{\rho}/a)^2}{4(1 + \tilde{\rho}/a)}. \tag{5}$$

We transform to the conventional dimensionless variables used in analyzing the motion of fast particles in matter, according to Ref. 8:

$$\varepsilon = E \frac{m_2 a_{TF}}{Z_1 Z_2 e^2 (m_1 + m_2)},$$

$$s_n(\varepsilon) = \frac{m_1 + m_2}{m_1} \frac{1}{4\pi a_{TF} Z_1 Z_2 e^2} S_n(E),$$

$$\omega(\varepsilon) = \frac{1}{\pi}\left(\frac{m_1 + m_2}{4 Z_1 Z_2 e^2 m_1}\right)^2 \Omega^2(E),$$

$a_{TF} = 0.8853 a_0/(Z_1^{2/3} + Z_2^{2/3})^{1/2}$ is the screening length in the Thomas–Fermi approximation, and a_0 is the Bohr radius.

From Eqs. (5) we obtain expressions for the ion stopping power $s_n(\varepsilon)$ and the straggling $\omega(\varepsilon)$ of the energy losses in the form

$$s_n(\varepsilon) = \left(\frac{a}{a_{TF}}\right)^2 \varepsilon f(\tilde{X}(\varepsilon)),$$

$$\omega(\varepsilon) = \left(\frac{a}{a_{TF}}\right)^2 \varepsilon^2 F(\tilde{X}(\varepsilon)),$$

$$\tilde{X}(\varepsilon) = \frac{1}{4}\frac{(a_{TF}/a)^2}{\varepsilon^2 + (a_{TF}/a)\varepsilon}. \tag{6}$$

It follows from Eqs. (6) that $s_n \to \ln \varepsilon/2\varepsilon$ and $\omega \to 1/4$ in the limit $\varepsilon \to \infty$, coinciding with the asymptotic behavior of the functions s_n and ω for the Coulomb interatomic interaction potential. In the limit $\varepsilon \to 0$ we have $s_n \to (a/a_{TF})^2 \varepsilon/2$ and $\omega \to (a/a_{TF})^2 \varepsilon^2/3$, consistent with the hard-sphere model.[4]

In low-energy experiments the ion stopping power is observed to have approximately a square-root dependence on the energy,[8] which differs from the dependence $s_n(\varepsilon)$ given by Eqs. (6). To bring Eqs. (6) closer to reality, it is proposed that the screening length a be assumed to depend on the ion energy. In this way the screening function in the interaction potential (1) is assumed to depend not only on the distance between the colliding particles, but also on the energy of the ion. If the dependence of the charge state of the ion (which largely dictates the screening function) in its motion through matter is taken into account,[8,9] the stated assumption appears entirely reasonable. To ensure that $s_n \to \sqrt{\varepsilon}$ in the limit $\varepsilon \to 0$, in Eqs. (6) we introduce the energy dependence of the screening length in the form

$$a = a_{TF}\varepsilon^{-1/4}/\beta, \tag{7}$$

where the parameter β can be regarded as a fitting parameter.

For this energy dependence of the screening length the stopping power and the straggling assume the form

$$s_n(\varepsilon) = \frac{\sqrt{\varepsilon}}{\beta^2} f(\tilde{X}(\varepsilon)),$$

$$\omega(\varepsilon) = \frac{\varepsilon^{3/2}}{\beta^2}(\tilde{X}(\varepsilon)),$$

$$\tilde{X}(\varepsilon) = \frac{\beta^2}{4(\varepsilon^{3/2} + \beta\varepsilon^{3/4})}. \tag{8}$$

In Figs. 1 and 2 the ion stopping powers and straggling of the energy losses, calculated from Eqs. (8), are compared with the corresponding functions for the Molière and Kr–C potentials given in Ref. 8. For $\beta = 0.54$ the ion stopping power obtained in this paper essentially coincides with the stopping power for the Kr–C potential at ion energies in the interval $10^{-3} \leqslant \varepsilon \leqslant 0.1$. For $\beta = 0.6$ these stopping powers essentially coincide in the energy interval $\varepsilon > 0.1$. The energy-loss straggling obtained in this paper for $\beta = 0.54$ is greater than the straggling for the Molière and Kr–C potentials. For $\beta = 0.7$ the straggling calculated from Eqs. (8) is close to the straggling for the Kr–C potential.
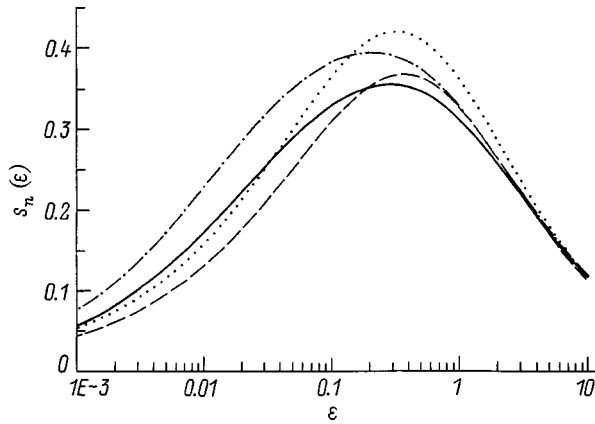
FIG. 1. Stopping powers of ions for various interaction potentials: the Kr–C potential (solid curve); the Molière potential (dot-dashed curve); calculated from Eq. (8) for $\beta=0.54$ (dotted curve); the same, for $\beta=0.6$ (dashed curve).

NUMERICAL SIMULATION OF THE RANGES OF IONS IN MATTER USING THE NEW MODEL POTENTIAL

We have developed a program for simulating the ranges of ions in matter by the Monte Carlo method, using the above-described model interaction potential. Let us first list the approximations used in the program. We consider a stream of ions with initial energy E_0 in normal incidence on the surface of a solid. The surface is located at the space coordinate $x=0$, and the solid occupies the half space $x>0$. As the ions pass through the material, they experience elastic collisions with target atoms with random variations of their energy and direction. We assume that in motion between successive elastic collision events an ion loses energy in inelastic collisions, which are treated in the continuous-slowing approximation. We trace the motion of the ion until its energy drops below a certain threshold level E_{th}. By analogy with Ref. 5, the threshold energy is defined in terms of the displacement energy of a target atom E_d. The computation is also stopped when the ion goes beyond the boundary
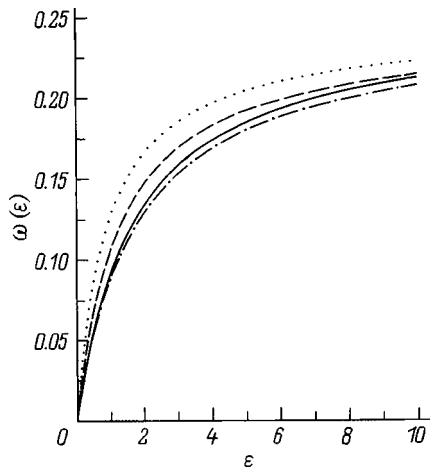

FIG. 2. Straggling of the energy losses for various interaction potentials: the Kr–C potential (solid curve); the Molière potential (dot-dashed curve); calculated from Eq. (8) for $\beta=0.54$ (dotted curve); the same, for $\beta=0.7$ (dashed curve).

of the solid ($x<0$). When an ion comes to rest in the material, the coordinate at which it stopped is determined and written to a file from which the spatial distribution of ions implanted in the material is constructed. The number of trials (number of ion trajectories traced) is set relatively high to ensure that the results will be statistically reliable.

In the Monte Carlo simulation of various processes random variables are customarily expressed in terms of one random variable ξ distributed uniformly in the interval from 0 to 1. In the simulation of ion ranges in matter the distance traversed by an ion between successive elastic collisions, the impact parameter, and the azimuthal scattering angle are regarded as random variables. The probability density function of the ion energy changing from ε to ε' as a result of inelastic slowing processes during its motion between two successive elastic collision events is determined in accordance with Ref. 7 from the relation

$$p(\varepsilon,\varepsilon')=\frac{1}{Q}\frac{\sigma(\varepsilon)}{\pi a_{TF}^2 s_e(\varepsilon)}\exp\left[-\int_{\varepsilon'}^{\varepsilon}\frac{1}{Q}\frac{\sigma(\varepsilon')}{\pi a_{TF}^2 s_e(\varepsilon')}d\varepsilon'\right],\tag{9}$$

where $Q=4m_1m_2/(m_1+m_2)^2$, $s_e(\varepsilon)$ is the stopping power of the ions in inelastic slowing, and $\sigma(\varepsilon)$ is the total elastic scattering cross section, which is defined in terms of the screening length, $\sigma(\varepsilon)=\pi a^2(\varepsilon)$.

On the basis of Eq. (7) we have $\sigma(\varepsilon)=\pi a_{TF}^2/(\beta^2\sqrt{\varepsilon})$. Equation (9) permits the energy ε' to be expressed in terms of the random variable ξ by means of the law of transformation of random variables.[10] Assuming that $s_e(\varepsilon)=k\sqrt{\varepsilon}$, we readily obtain

$$\varepsilon'=\varepsilon^{kQ\beta^2}.\tag{10}$$

The inelastic energy losses of an ion is described in the continuous-slowing approximation by the equation[7]

$$\frac{d\varepsilon}{dl}=-As_e(\varepsilon),$$

where $A=Qn\pi a_{TF}^2$, and n is the density of target atoms.

The determination of the distance traversed by an ion between elastic collisions from the above equation is obvious:

$$l=\frac{2}{Ak}(\sqrt{\varepsilon}-\sqrt{\varepsilon'}).\tag{11}$$

Substituting Eq. (10) into (11), we express the distance l in terms of the random variable ξ:

$$l=\frac{2\sqrt{\varepsilon}}{Ak}(1-\xi^{kQ\beta^2/2}).\tag{11a}$$

The value $\xi=0$ corresponds to the maximum distance between collisions $l=2\sqrt{\varepsilon}/(Ak)$. For $\xi=1$ we have $l=0$. If inelastic processes can be ignored, $k\rightarrow0$, it follows from Eq. (11a) that $l\rightarrow\ln(1/\xi)/n\sigma(\varepsilon)$. This condition corresponds to the limiting case of free motion of the ion between elastic collisions.[1]

The variation of the ion parameters as a result of elastic scattering by a target atom are described by relations taken from Ref. 1:

$$\cos \Theta = \left(1 + \frac{m_2}{m_1}\cos \chi\right) \Big/ \sqrt{1 + 2\frac{m_2}{m_1}\cos \chi + \left(\frac{m_2}{m_1}\right)^2},$$

$$\frac{\varepsilon''}{\varepsilon'} = \left(1 + 2\frac{m_2}{m_1}\cos \chi + \left(\frac{m_2}{m_1}\right)^2\right) \Big/ \left(1 + \frac{m_2}{m_1}\right)^2,$$

where Θ is the scattering angle of the ion in the laboratory frame, ε' is the energy of the ion before collision, ε'' is its energy after collision, and $\cos\chi$ is defined by Eq. (4), in which t is proportional to the square of the impact parameter and is modeled by the random variable according to the relation $t = \xi$.

The coordinate x_k of the kth collision of an ion in the material is determined by summing the corresponding displacements of the ion along the x axis as it moves between successive elastic collisions:

$$x_k = \sum_{j=0}^{k-1} l_j \mu_j,$$

where l_j is the distance traversed by the ion between the jth and the $(j+1)$st elastic collision, and μ_j is the cosine of the angle between the trajectory of the ion after the jth collision and the x axis.

According to Ref. 5, μ_j obeys the recursion relation

$$\mu_j = \mu_{j-1}\cos \Theta + \sqrt{1 - \mu_{j-1}^2}\sin \Theta \cos \psi.$$

The azimuthal scattering angle ψ is expressed in terms of the random variable ξ by the relation $\psi = 2\pi\xi$.

Our tracing of the trajectories of N ions from their first entry into the material until they stop produces a set of values of the coordinates at which the ions stop in the material, $\hat{x}_i$, $i = 1, \ldots, N$. These calculations are used to plot histograms of the spatial distribution of implanted ions and to determine certain averages: the average projected range R_p and the rms deviation ΔR_p of the projected ranges:

$$R_p = \frac{1}{N}\sum_{i=1}^{N} \hat{x}_i, \quad \Delta R_p = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \hat{x}_i^2 - R_p^2}. \quad (12)$$

RESULTS OF MONTE CARLO CALCULATIONS

To describe the ranges of ions in matter by the given model, it is necessary to specify k, which characterizes the inelastic energy losses, and β, which characterizes the screening length. We regard β and k as fitting parameters and determine them by trial and error in comparing the results of the numerical calculations with experiment. In Table I the experimental values of R_p and ΔR_p from Ref. 3 are compared with the results of the Monte Carlo calculations. For each energy the calculations are carried out by tracking 10^5 histories. The relative error of calculation of the average value of R_p, estimated in accordance with Ref. 10, does not exceed 0.3% in this case. The values of β and k for the given ion–target pairs are determined in the stage of preliminary calculations. We have $\beta = 0.67$ and $k = 0.04$ for the pair Cu–C, $\beta = 0.6$ and $k = 0.04$ for Rb–C, and $\beta = 0.57$ and $k = 0.05$ for Rb–B. It is evident from the table that the results of the numerical calculations are in good agreement with the

TABLE I. Parameters of the experimental and theoretical ion ranges.

Ion	Target	E, keV	ε	Experiment		Theory	
				R_p, Å	ΔR_p, Å	R_p, Å	ΔR_p, Å
Cu	C	30	0.250	280	90	300	88
		50	0.416	430	130	441	124
		79	0.582	570	160	579	157
		100	0.832	785	215	791	205
		150	1.248	1180	320	1162	283
		200	1.664	1547	400	1555	362
Rb	C	10	0.047	104	40	121	34
		30	0.143	210	70	242	65
		50	0.238	330	90	347	90
		80	0.380	500	145	496	123
		100	0.475	590	160	593	144
		150	0.713	850	215	846	195
		200	0.951	1077	270	1106	245
Rb	B	20	0.105	170	45	167	43
		50	0.264	325	80	317	78
		100	0.527	565	150	552	125
		300	1.581	1550	320	1592	310

experimental data. The maximum relative deviation of the theoretical values of R_p and ΔR_p from the experimental does not exceed 17%. The absolute deviation does not exceed 42 Å. Figure 3 shows histograms of the spatial distribution of Cu ions in a carbon target, calculated by the Monte Carlo method, for two approximations: with inelastic losses taken into account ($k = 0.04$) and without inelastic losses ($k = 0$). When inelastic losses are disregarded, the maximum of the spatial distribution shifts toward higher values of x, and the distribution broadens. For example, as k is varied from 0.04 to 0, R_p increases from 1555 Å to 1742 Å, and ΔR_p increases from 362 Å to 440 Å. It follows from Fig. 3 that the spatial distribution of implanted 200-keV Cu ions in a carbon target is very accurately described by a Gaussian distribution:

$$Ni(x) = \frac{1}{\Delta R_p \sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{x - R_p}{\Delta R_p}\right)^2\right).$$
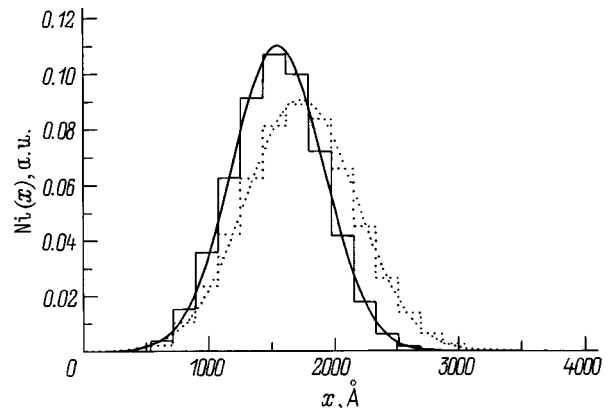

FIG. 3. Spatial distribution of implanted Cu ions with energy $E = 200$ keV in a carbon target. The histograms are calculated by the Monte Carlo method for $\beta = 0.67$, $k = 0.04$ (solid curve) and $k = 0$ (dotted curve). The curves represent Gaussian distribution functions for $R_p = 1555$ Å, $\Delta R_p = 362$ Å (solid curve) and for $R_p = 1742$ Å, $\Delta R_p = 440$ Å (dotted curve).
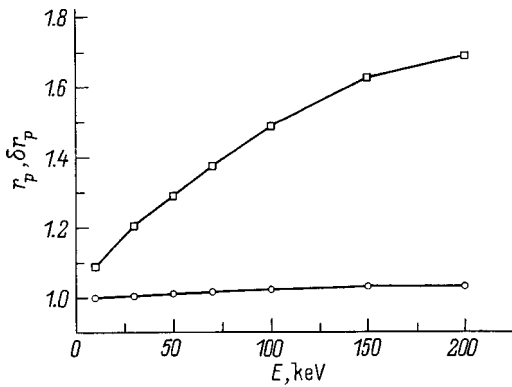
FIG. 4. Ratio of the ranges of Cu ions in a carbon target, calculated by the Monte Carlo method in the modified hard-sphere approximation and with our new model potential, $\beta=0.67$, $k=0.04$; the circles represent r_p, and the squares represent δr_p.

These results show that the proposed interaction potential can be used to described the ranges of ions in matter with high accuracy when the stopping of the ions is governed by both elastic and inelastic energy losses. We can use the newly developed program to determine the domain of applicability of the modified hard-sphere model proposed in Refs. 4 and 5 for calculating the ranges of low-energy ions ($\varepsilon < 0.1$) in amorphous substances. In the modified hard-sphere model scattering is isotropic in the CM frame, and the total elastic scattering cross section of the ions is described, according to Ref. 4, by the relation $\sigma(\varepsilon)=2\pi a_{TF}^2 s_n(\varepsilon)/\varepsilon$. In the Monte Carlo calculations we determine $s_n(\varepsilon)$ for the modified hard-sphere mode from Eq. (8).

Corrections corresponding to this model have been introduced in the program for the Monte Carlo calculations of the ion ranges. The results of the calculations for the pair Cu–C in the modified hard-sphere model approximation are shown in relative form in Fig. 4. The quantities r_p and δr_p plotted in the figure are the ratio of the corresponding quantities R_p and ΔR_p calculated in the modified hard-sphere model approximation to the ranges calculated in the approximation of the new (real) interaction potential. It follows from Fig. 4 that the projected ion range R_p calculated in the hard-sphere model approximation essentially coincide with the range calculated for the real interaction potential over the entire energy range. The rms deviation ΔR_p in the modified hard-sphere model approximation is too high. The relative function δr_p increases as the energy is increased. We find that $\delta r_p < 1.1$ at ion energies $E \leqslant 10$ keV ($\varepsilon \leqslant 0.083$), and $\delta r_p \approx 1.2$ at an ion energy $E = 30$ keV ($\varepsilon = 0.25$). This comparison shows, therefore, that the modified hard-phase model can be used to calculate R_p and ΔR_p in the energy range $\varepsilon < 0.2$. In the range of higher energies the modified hard-sphere model can be used to calculate the projected range R_p, but the rms deviation ΔR_p is much too high in this case.

CONCLUSIONS

We have proposed a new model interaction potential in the form of a screened Coulomb potential. We have derived analytical expressions for the scattering angle in the center-of-mass frame as a function of the impact parameter and the ion energy. We have obtained analytical expressions for the stopping power of ions in elastic collisions and the straggling of their energy losses. We have written a program for simulating the ranges of ions in matter by the Monte Carlo method, taking elastic energy losses into account in the approximation of the new interaction potential and taking inelastic losses into account in the continuous-slowing approximation. We have calculated the ranges R_p ΔR_p for Cu and Rb ions in carbon and boron targets. The results of calculations agree very accurately with the experimental results, indicating that we have made a good choice of interaction potential. Our comparative calculations of the ranges of ions in matter in the approximation of the new model interaction potential and in the approximation of the previously developed modified hard-sphere model have shown that the modified hard-sphere model can be used to calculate R_p and ΔR_p at low energies $\varepsilon < 0.2$. The projected ion range R_p in the modified hard-sphere model essentially coincides over the entire energy range with the projected range calculated in the approximation of the new, more realistic interaction potential. This agreement can be utilized in the future to determine the fitting parameter β in the equation for the screening length of the new model potential by a quick and easy procedure, which does not use Monte Carlo calculations and relies instead on the simpler technique of comparing the experimental values of R_p with analytical results[7] obtained in the modified hard-sphere model.

[1] W. Eckstein, *Computer Simulation of Ion-Solid Interactions* (Springer-Verlag, Berlin-New York, 1991).

[2] A. F. Akkerman, *Modeling of Charged Particle Trajectories in Matter* (Énergoatomizdat, Moscow, 1991), 201 pp.

[3] M. Grande, F. C. Zawislak, D. Fink, and M. Behar, Nucl. Instrum. Methods Phys. Res. B **61**, 282 (1991).

[4] E. G. Sheĭkin, Zh. Tekh. Fiz. **66**(10), 63 (1996) [Tech. Phys. **41**, 1005 (1996)].

[5] E. G. Sheĭkin, Zh. Tekh. Fiz. **67**(10), 16 (1997) [Tech. Phys. **42**, 1128 (1997)].

[6] E. G. Sheĭkin, Zh. Tekh. Fiz. **68**(9), 33 (1998) [Tech. Phys. **43**, 1039 (1998)].

[7] E. G. Sheĭkin, Zh. Tekh. Fiz. **69**(2), 93 (1998) [Tech. Phys. **44**, 218 (1999)].

[8] M. A. Kumakhov and F. F. Komarov, *Energy Losses and Ranges of Ions in Solids* [in Russian], Minsk (1979), 320 pp.

[9] Y.-H. Ohtsuki, *Charged Beam Interaction with Solids* [Taylor and Francis, London (1983); Mir, Moscow (1985), 280 pp.].

[10] I. M. Sobol', *Monte Carlo Numerical Methods* [in Russian], Nauka, Moscow (1973), 312 pp.

Translated by James S. Wood

Instability of a charged spherical drop moving relative to a medium

A. I. Grigor'ev, V. A. Koromyslov, and S. O. Shiryaeva

Yaroslavl State University, 150000 Yaroslavl, Russia

It is shown that the critical self-charge for the onset of instability of a charged drop in a flow of
an ideal fluid decreases as the flow velocity of the fluid past the drop increases, i.e., a
complex instability arises which is a superposition of the instabilities of the free surface of the
drop with respect to the tangential discontinuity of the velocity field at the free surface of
the drop and with respect to the self-charge. © *1999 American Institute of Physics.*
[S1063-7842(99)00205-6]

Charged particles moving relative to a medium are encountered in various problems of technical physics, geophysics, and technology.[1–3] However, despite a great number of studies of the breakup of freely falling drops in the atmosphere (see the survey in Ref. 4), problems associated with the functional relations involved in the development of instability in drops with respect to their self-charge and the tangential discontinuity of the velocity field at the free surface of the drop have not been investigated to date. This problem has very important bearing on thunderstorm electricity in regard to the study of the physical mechanism underlying the initiation of lightning discharges. According to Refs. 5–7, lightning can be initiated from a corona discharge in the vicinity of a freely falling, large, waterlogged hailstone. However, the strength of the external electric field required to trigger a corona discharge in the vicinity of a hailstone is far above the levels actually observed in thunderclouds. On the other hand, it is known from experiments[6] that airflow around the boundary lowers the critical external electric field for triggering a corona discharge. This kind of phenomenon can be attributed to the interaction of two types of instability of the free surface of the water layer on the boundary surface: its instability with respect to the induced charge,[3] and its instability with respect to the tangential discontinuity of the velocity field.[8]

Inasmuch as the critical conditions for instability of capillary waves on a charged surface of a liquid in the presence of a tangential discontinuity of the velocity field does not depend on the viscosity,[9] the discussion that follows can be simplified by confining it to the ideal-fluid approximation.

1. Let an ideal, incompressible, dielectric medium of density ρ_1 and dielectric constant ε move with a constant velocity $\mathbf{U}$ relative to a spherical drop of radius R of an ideal, perfectly conducting fluid of density ρ_2, which carries a charge Q. We wish to determine the critical conditions for instability of capillary oscillations of the drop under the conditions described here.

The entire investigation is carried out in spherical coordinates with origin at the center of the drop in the linear approximation with respect to the perturbation $\xi(\Theta,t)$ of the equilibrium spherical surface of the drop due to thermal capillary waves of amplitude $\sim 10^{-8}$ cm. We write the equation for the perturbed surface of the drop in the form $r(\Theta,t) = R + \xi(\Theta,t)$. We assume that the wave motions in the surrounding medium and in the drop, like the potential Φ of the electric field of the self-charge in the vicinity of the drop, are of the potential type with harmonic velocity potentials Ψ_1 and Ψ_2, respectively, which obey the Laplace equation. The problem is stated mathematically in the form

$$\Delta\Phi=0; \quad \Delta\Psi_i=0 \quad (i=1;2); \tag{1}$$

$$r=0: \quad \nabla\Psi_2=0; \quad \Phi=\text{const}; \tag{2}$$

$$r=R+\xi: \quad \Phi=\text{const};$$

$$r=R: \quad \frac{\partial\xi}{\partial t}+\frac{1}{r^2}\frac{\partial\Psi_1}{\partial\Theta}\frac{\partial\xi}{\partial\Theta}=\frac{\partial\Psi_1}{\partial r}; \tag{3}$$

$$\frac{\partial\xi}{\partial t}=\frac{\partial\Psi_2}{\partial r}; \tag{4}$$

$$-\rho_1\frac{\partial\Psi_1}{\partial t}+\rho_2\frac{\partial\Psi_2}{\partial t}-\rho_1(\nabla\Psi_1)^2-P_E+P_\sigma=0; \tag{5}$$

$$r\to\infty: \quad \nabla\Psi_1=\mathbf{U}; \quad \Phi\to 0. \tag{6}$$

The squared term $\sim(\nabla\Psi_1)^2$ is retained in Eq. (5) because it contains a term of zeroth-order smallness; P_σ is the pressure due to surface tension, P_E is the pressure created by the electric field on the free surface of the drop, and σ is the coefficient of surface tension.

2. To simplify the solution of the problem, it is useful to transform to dimensionless variables, in which $R=1$, $\sigma=1$, and $\rho_2=1$. All other quantities (for which we keep the same notation as before) are then expressed in units of their characteristic values

$$r_*=R; \quad t_*=R^{3/2}\rho_2^{1/2}\sigma^{-1/2}; \quad U_*=R^{-1/2}\rho_2^{-1/2}\sigma^{1/2};$$

$$p_*=R^{-1}\sigma; \quad Q_*=R^{3/2}\sigma^{1/2}.$$

Here, as in the similar Helmholtz instability problem,[8,9] we set $\rho_1/\rho_2\equiv\rho$. We seek a solution of Eqs. (1), together with Eqs. (2) and (6), for Ψ_1 and Ψ_2 and for the perturbation $\xi(\theta,t)$ in the form

$$\Psi_1(\mathbf{r},t)=\varphi+\sum_n A_n r^{-(n+1)}Y_n(\mu)\exp(St);$$

$$\Psi_2(\mathbf{r},t)=\sum_n B_n r^n Y_n(\mu)\exp(St); \qquad (7)$$

$$\xi(\mathbf{r},t)=\sum_n Z_n Y_n(\mu)\exp(St);$$

where $Y_n(\mu)$ are normalized spherical harmonics, $\mu\equiv\cos\theta$, S is the complex frequency, and A_n, B_n, and Z_n are quantities of the same order of smallness; φ is the velocity potential of the fluid flow in the vicinity of the unperturbed spherical surface, the gradient of which gives an expression for the freestream velocity field:

$$\mathbf{V}=\nabla\varphi=-\frac{R^3}{2r^3}[3\mathbf{n}(\mathbf{U}\cdot\mathbf{n})-\mathbf{U}]+\mathbf{U},$$

where $\mathbf{n}$ is the vector normal to the surface of the drop.

We assume for definiteness that the velocity $\mathbf{U}$ is in the same direction as the z axis in a Cartesian coordinate system. Making use of the fact that $\mathbf{e}_z\equiv\mathbf{e}_r\cos\Theta-\mathbf{e}_\Theta\sin\Theta$ and $\mathbf{n}\equiv\mathbf{e}_r$, we write an expression for the fluid flow velocity in the vicinity of the drop in spherical coordinates:

$$\mathbf{V}=U\cos\Theta\left[1-\frac{R^3}{r^3}\right]\mathbf{e}_r-U\sin\Theta\left[1+\frac{R^3}{2r^3}\right]\mathbf{e}_\Theta. \qquad (8)$$

Substituting the solutions (7) and (8) into the boundary conditions (3)–(5) and invoking the well-known[10] recursion relation

$$\sin\Theta\frac{\partial Y_n}{\partial\theta}=\frac{n(n+1)}{\sqrt{(2n+1)(2n+3)}}Y_{n+1}$$
$$-\frac{n(n+1)}{\sqrt{(2n+1)(2n-1)}}Y_{n-1},$$

we obtain the following in the linear approximation with respect to the small parameters:

$$r=1: \quad \sum_n SZ_n Y_n(\mu)\exp(St)-\frac{3}{2}U\sum_n Z_n$$

$$\times\left(\frac{n(n+1)}{\sqrt{(2n+1)(2n+3)}}Y_{n+1}-\frac{n(n+1)}{\sqrt{(2n+1)(2n-1)}}Y_{n-1}\right)$$

$$\times\exp(St)+\sum_n (n+1)A_n Y_n(\mu)\exp(St)=0; \qquad (9)$$

$$\sum_n SZ_n Y_n(\mu)\exp(St)=\sum_n nB_n r^n Y_n(\mu)\exp(St); \qquad (10)$$

$$-\rho S\sum_n A_n Y_n\exp(St)+S\sum_n B_n Y_n$$

$$\times\exp(St)+3\rho U\sum_n A_n\left[\frac{n(n+1)}{\sqrt{(2n+1)(2n+3)}}Y_{n+1}\right.$$

$$\left.-\frac{n(n+1)}{\sqrt{(2n+1)(2n-1)}}Y_{n-1}\right]$$

$$\times\exp(St)-\frac{Q^2}{4\pi\varepsilon}\sum_n (n-1)Z_n Y_n$$

$$\times\exp(St)+\sum_n (n-1)(n+2)Z_n Y_n\exp(St)=0, \qquad (11)$$

where the relations

$$r=1: \quad \nabla\Psi_1=\left[-\sum_n A_n(n+1)Y_n(\mu)\exp(St)\right]\mathbf{e}_r$$

$$+\left[-\frac{3}{2}U\sin\Theta+\sum_n A_n\frac{\partial Y_n}{\partial\Theta}\exp(St)\right]\mathbf{e}_\Theta;$$

$$(\nabla\Psi_1)^2\approx\frac{9}{4}U^2\sin^2\Theta-3U\sum_n A_n$$

$$\times\left[\frac{n(n+1)}{\sqrt{(2n+1)(2n+3)}}Y_{n+1}\right.$$

$$\left.-\frac{n(n+1)}{\sqrt{(2n+1)(2n-1)}}Y_{n-1}\right]\exp(St);$$

have been taken into account, and the expressions for the pressures $P_\sigma(\xi)$ and $P_E(\xi)$ are written in the form[9]

$$P_\sigma(\xi)=\sum_n (n-1)(n+2)z_n Y_n(\mu)\exp(St);$$

$$P_E=-\frac{Q^2}{4\pi\varepsilon}\sum_n (n-1)Z_n Y_n(\mu)\exp(St).$$

Multiplying Eqs. (9)–(11) by Y_m and integrating over the angle Θ, we obtain a system of homogeneous equations for the unknown amplitudes A_n, B_n, and Z_n:

$$SZ_n-\frac{3}{2}U(\alpha_n Z_{n-1}-\beta_n Z_{n+1})+(n+1)A_n=0; \qquad (12)$$

$$SZ_n=nB_n; \qquad (13)$$

$$-\rho SA_n+SB_n+3\rho U[\alpha_n A_{n-1}-\beta_n A_{n+1}]$$

$$-\left[\frac{Q^2}{4\pi\varepsilon}(n-1)-(n-1)(n+2)\right]Z_n=0;$$

$$\alpha_n\equiv\frac{n(n-1)}{\sqrt{(2n-1)(2n+1)}}; \quad \beta_n\equiv\frac{(n+1)(n+2)}{\sqrt{(2n+1)(2n+3)}}. \qquad (14)$$

From Eqs. (12) and (13) we can readily obtain

$$A_n = \frac{1}{(n+1)} \left\{ \frac{3}{2} U(\alpha_n Z_{n-1} - \beta_n Z_{n+1}) - SZ_n \right\};$$

$$B_n = \frac{1}{n} SZ_n.$$

Substituting these relations into Eq. (14), we obtain an infinite system of homogeneous algebraic equations that can be used to find the amplitudes of the capillary oscillations of the drop:

$$\rho U^2 K_n Z_{n-2} - \rho USL_n Z_{n-1} + \{ \varkappa_n S^2 - \rho U^2 M_n + \gamma_n \} Z_n$$
$$+ \rho USL_n Z_{n+1} + \rho U^2 J_n Z_{n+2} = 0;$$

$$W \equiv \frac{Q^2}{4\pi\varepsilon}; \quad M \equiv \frac{9\alpha_n\beta_{n-1}}{2n} + \frac{9\beta_n\alpha_{n+1}}{2(n+2)}; \quad K_n \equiv \frac{9\alpha_n\alpha_{n-1}}{2n};$$

$$L_n \equiv \frac{(9n+6)\alpha_n}{2n(n+1)}; \quad I_n \equiv \frac{(9n+12)\beta_n}{2(n+1)}; \quad J_n \equiv \frac{9\beta_n\beta_{n+1}}{2(n+2)};$$

$$\varkappa_n = \equiv \frac{\rho}{(n+1)} + \frac{1}{n}; \quad \gamma_n \equiv (n-1)[n+2-W]. \qquad (15)$$

3. The necessary and sufficient condition for the existence of a solution of the homogeneous system (15) is that the determinant constructed from the coefficients of the unknown amplitudes Z_n be equal to zero [see Eq. (16)].

This equation is the dispersion relation governing the spectrum of capillary oscillations of the drop as a function of dimensionless physical parameters W, U, and ρ. The spectrum changes as these quantities are varied, and for definite values of W, U, and ρ certain solutions S_n^2 can pass through zero and become positive. When this condition is satisfied, the amplitudes of the corresponding capillary oscillation grow exponentially with time, i.e., the drop becomes unstable[3] and decays according to the law described in Ref. 11. The condition for the onset of zero solutions of the dispersion relation is that the free coefficient of the dispersion-relation (16) be equal to zero [see Eq. (17)]:

$$\begin{vmatrix} \varkappa_2 S^2 - \rho U^2 M_2 + \gamma_2 & \rho USL_2 & \rho U^2 J_2 & 0 & \ldots \\ -\rho USL_3 & \varkappa_3 S^2 - \rho U^2 M_3 + \gamma_3 & \rho USL_3 & \rho U^2 J_3 & \ldots \\ \rho U^2 K_4 & -\rho USL_4 & \varkappa_4 S^2 - \rho U^2 M_4 + \gamma_4 & \rho USL_4 & \ldots \\ 0 & \rho U^2 K_5 & -\rho USL_5 & \varkappa_5 S^2 - \rho U^2 M_5 + \gamma_5 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{vmatrix} = 0. \qquad (16)$$

$$\begin{vmatrix} \gamma_2 - \rho U^2 M_2 & 0 & \rho U^2 J_2 & 0 & 0 & \ldots \\ 0 & \gamma_3 - \rho U^2 M_3 & 0 & \rho U^2 J_3 & 0 & \ldots \\ \rho U^2 K_4 & 0 & \gamma_4 - \rho U^2 M_4 & 0 & \rho U^2 J_3 & \ldots \\ 0 & \rho U^2 K_5 & 0 & \gamma_5 - \rho U^2 M_5 & 0 & \ldots \\ 0 & 0 & \rho U^2 K_6 & 0 & \gamma_6 - \rho U^2 M_6 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{vmatrix} = 0. \qquad (17)$$

The system (15) is infinite, and the determinant (17) is also of infinite order. It is convenient to seek the critical conditions for the onset of instability of the capillary oscillations (i.e., the critical interdependence of W, U, and ρ) by successive approximations, considering two, three, four, and more equations of the system (15). On the other hand, certain conclusions can be drawn at once about the solutions of Eq. (17).

If mode interaction is disregarded, the critical conditions for instability of the nth capillary mode have the simple analytical form

$$(n+2) - \frac{1}{(n-1)} \rho U^2 M_n - W = 0. \qquad (18)$$

It is evident from this result that the necessary value of the Rayleigh parameter W for the onset of instability of the drop with respect to the self-charge decreases as the velocity and density of the medium increase. For $W=0$ we obtain the condition for the onset of instability of the drop with respect to the tangential discontinuity of the velocity field:

$$\rho U^2 \geqslant (n-1)(n+2) \frac{1}{M_n}.$$

It is evident from Eq. (18) that the critical value of the Rayleigh parameter $W = W_*$ for the onset of instability, evaluated for $n=2$, decreases rapidly as the flow velocity around the drop increases:

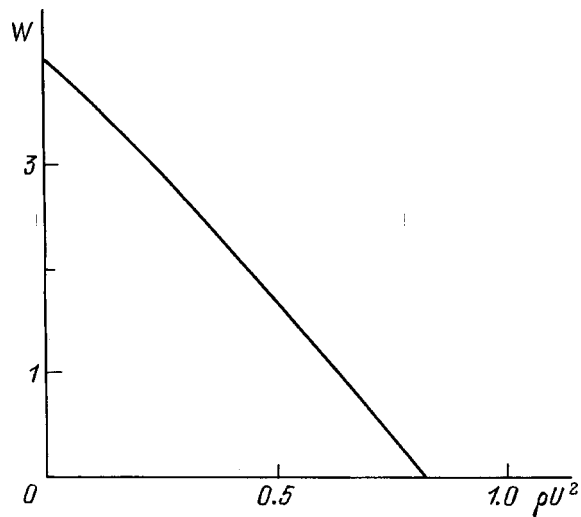$$W_* = 4 - \rho U^2 M_2. \qquad (19)$$

FIG. 1. Dependence of the critical Rayleigh parameter W on the dimensionless parameter ρU^2 for the fundamental (second) mode.
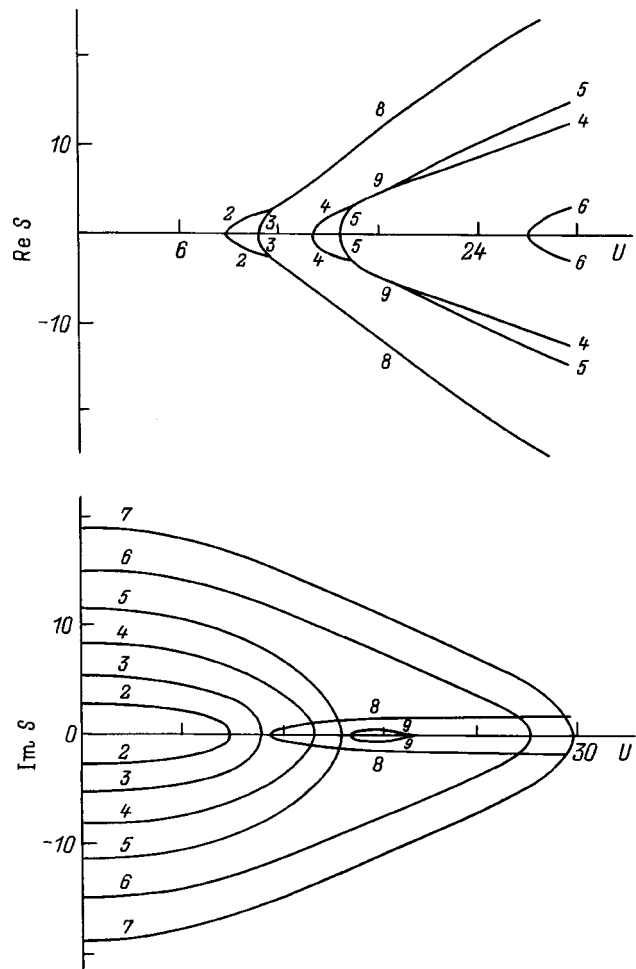

FIG. 3. Dependence of the real and imaginary components of the frequency S on the flow velocity U.
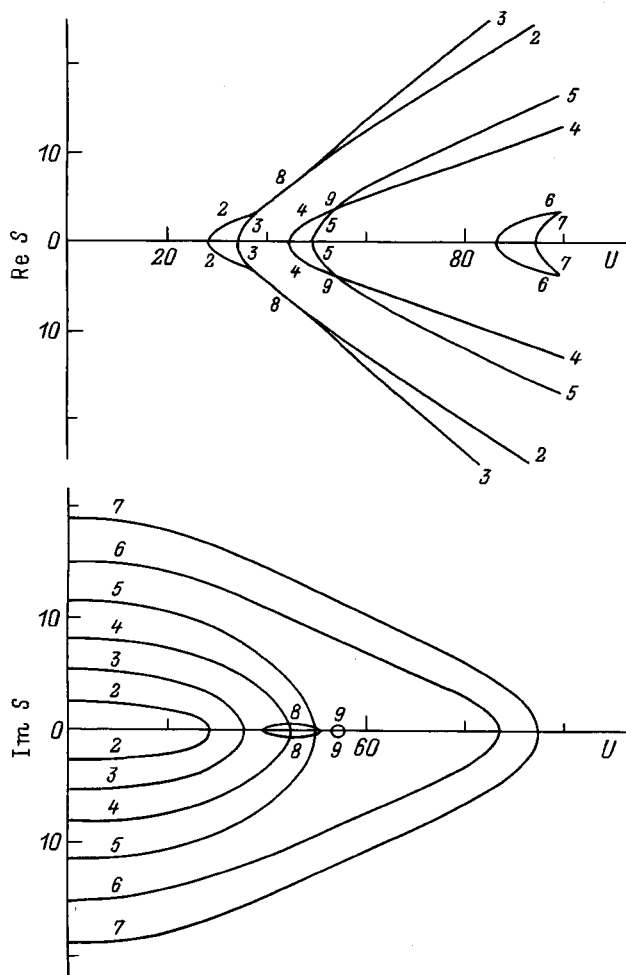

FIG. 2. Dependence of the real and imaginary components of the dimensionless frequency S of capillary motions of the fluid on the dimensionless velocity U of the flow past the drop.

The critical relation $W = W(\rho U^2)$, calculated numerically from (17) for $n = 2$ using the first six equations of the system (15), is plotted in Fig. 1. A comparison of the results of the numerical analysis with the analytical relation (19) show that the solutions converge very rapidly — in two or three steps — in the given successive-approximation procedure. It is evident from Fig. 1 that the relation $W = W(\rho U^2)$ differs very little from the linear relation predicted by Eq. (18).

From the geophysical standpoint the rapid decrease in the critical value of the parameter W for the onset of intrinsic-charge instability of the drop as ρU^2 increases has motivated renewed efforts to construct a physical model of the initiation of lightning discharges.[6,7,12] Existing notions are based on the concept of a corona discharge being triggered in the vicinity of a large thawing hailstone falling freely in a thundercloud, and the results obtained above canbe used to reconcile the initiation model with the realities of a thundercloud (from the measured charges at the boundaries, the strength of the electric field in the interior of the cloud, and the rate of descent of the hailstones).

The dispersion relation of the problem without mode interaction also has a simple form:
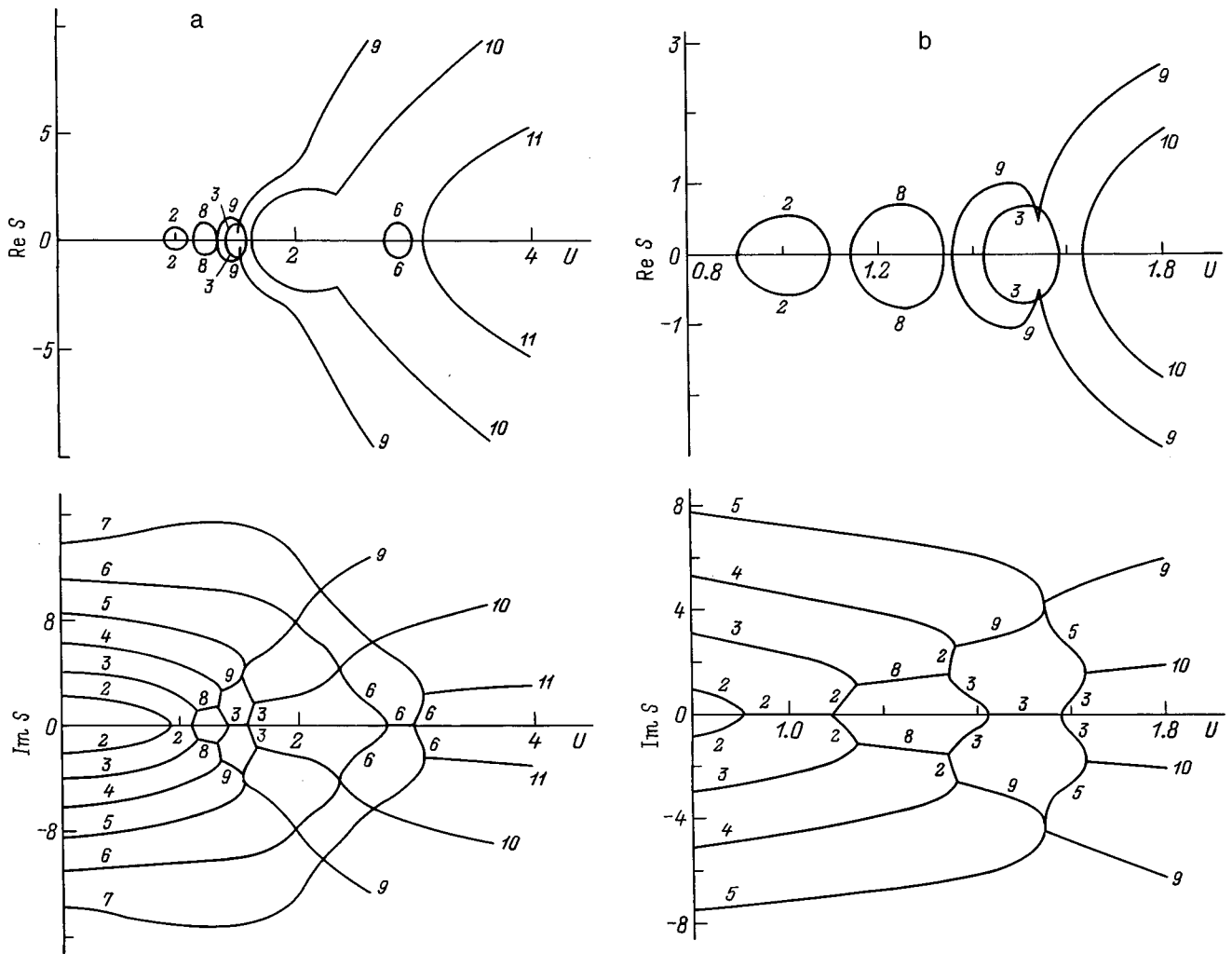
FIG. 4. a: Dependence of the real and imaginary components of the frequency S on the velocity U, $\rho=1$, $W=0$ (the mode orders are indicated by the curve numbers 2–7); b: dependence analogous to Fig. 4a, drawn to a larger scale.

$$S_n^2=[n(n-1)[W-n+2]-n\rho U^2 M_n]\left(\frac{n\rho}{(n+1)}+1\right)^{-1} \quad (20)$$

and in the limit $U\rightarrow0$ coincides with the well-known dispersion relation for the capillary oscillations of a charged drop in a dielectric medium.[9]

4. Numerical calculations by successive approximations according to the dispersion relation (16) show that the qualitative behavior of the capillary frequency of the drop as a function of the freestream velocity $S=S(U)$ is the same for the different modes at $W=0$; it is shown in Fig. 2 for the first six modes when $\rho=10^{-3}$. The curves are labeled by the mode orders 2–7. One noticeable phenomenon is the pairwise interaction of even and odd modes, which generates the oscillatory solutions 8 and 9. Branches 8 and 9 of the dispersion relation, which differ from branches 2–7 in that they have both real and imaginary components, are the result of interaction between the second and third modes and between the fourth and fifth modes, respectively. The parts of the $\mathrm{Re}\,S=\mathrm{Re}\,S(U)$ curves situated in the region $\mathrm{Re}\,S>0$ determine the instability growth rates of the corresponding capillary modes of the drop. In a certain velocity range, therefore,

the instability of the drop is oscillatory (branches 8 and 9), as is typical of instability of the Kelvin–Helmholtz type.[8,9–13]

Another striking feature is that to the right of the velocity interval in which the second and third modes interact, the growth rate of the aperiodic instability of the third mode is higher than the growth rate of the aperiodic instability of the second mode. A similar pattern is observed for the fourth and fifth modes. This result can be interpreted in the light of Eqs. (20), which give the dispersion relation without mode interaction. It is evident at once that the derivative of S_n^2 with respect to ρU^2, which gives the growth rate of S_n as U varies, is proportional to $n\cdot M_n\sim n^3$, i.e., without mode interaction the instability growth rates of the higher modes increase with the mode order as the surrounding flow accelerates. An increase in the wave number of the most unstable mode as the shear flow velocity increases has also been noted in an analysis of the classical Kelvin–Helmholtz instability on the flat free surface of a fluid.[9]

The indicated difference in the growth rates of the second and third modes should be mirrored phenomenologically in the relations governing the breakup of an unstable drop. For example, if the maximum instability growth rate occurs
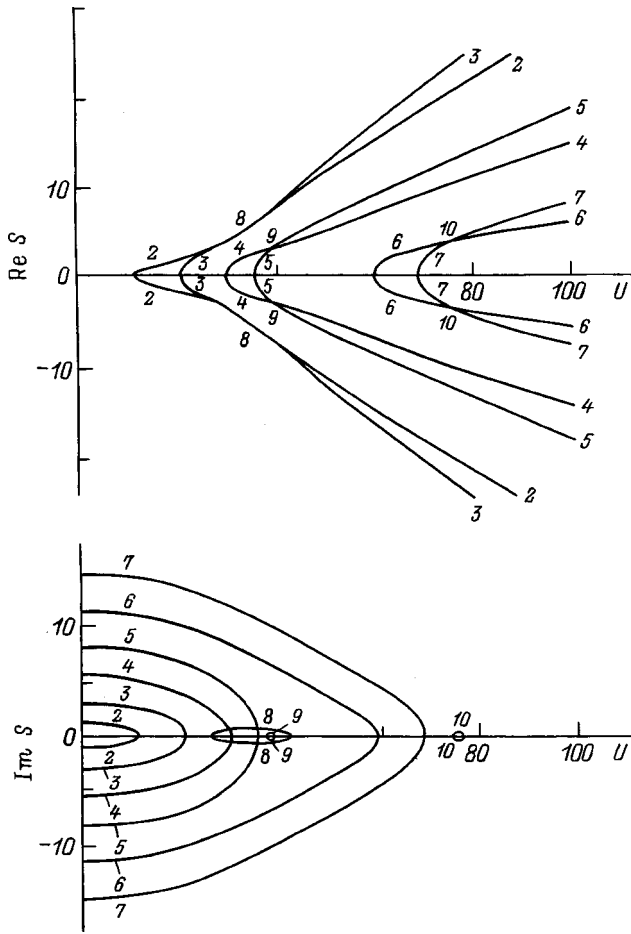
FIG. 5. Dependence of the real and imaginary components of the frequency S on the velocity U for $\rho = 10^{-3}$ and a Rayleigh-subcritical charge $W = 3.5$. Curve *10* is the result of interaction between the sixth and seventh modes.
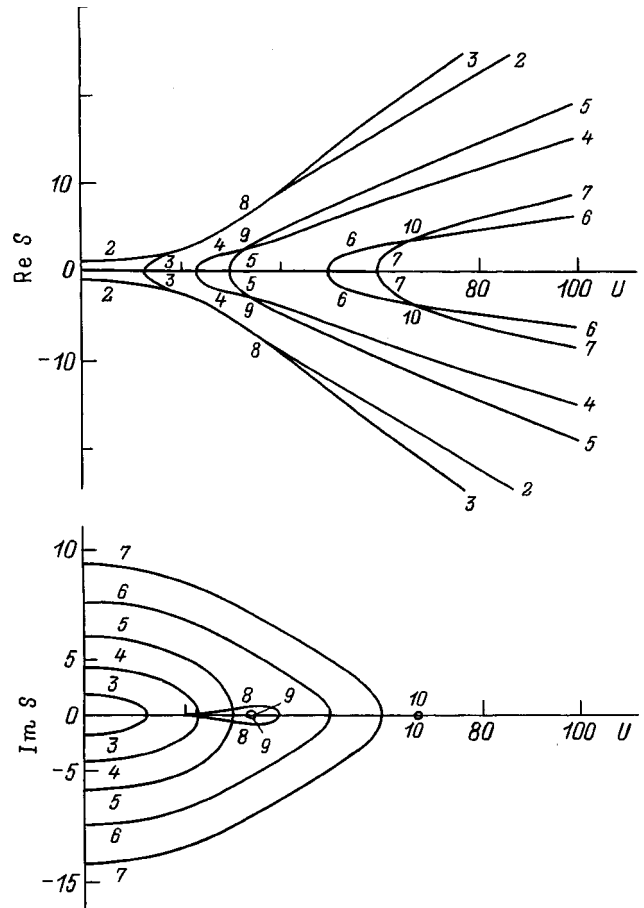

FIG. 6. Dependence of the real and imaginary components of the frequency S on the velocity U for $\rho = 10^{-3}$ and a Rayleigh-supercritical charge $W = 4.5$.

for the second (fundamental) mode for given densities of the drop and the medium and for a given passing flow velocity (the rate of descent of the drop in a nonmoving medium), the unstable drop acquires a shape that resembles a spheroid of revolution [characterized by the second Legendre polynomial $P_2(\mu)$] and then breaks up into two approximately equal daughter droplets. If the maximum instability growth rate occurs for the third mode, the shape acquired by the drop resembles a parachute [characterized by the third Legendre polynomial $P_3(\mu)$], and it breaks up into a set of small and somewhat larger droplets. These two possible breakup channels have been observed experimentally.[4] The oscillatory instability of a freely falling drop differs in experiments.[4]

It should also be noted that the critical dependences on W, ρ, and U for the sixth and seventh modes have been obtained for a low degree of mode interaction, since the determinant for the first six equations of the system (15) was used in the numerical calculations. This consideration accounts for a certain difference of the qualitative behavior of the $\mathrm{Im}\, S = \mathrm{Im}\, S(U)$ curves for the sixth and seventh modes from the lower modes.

The region of oscillatory instability becomes wider as the density of the medium increases. This situation is illustrated for $\rho = 10^{-2}$ in Fig. 3. The curves are numbered the same as in Fig. 2. Clearly, not only does the region of

Kelvin–Helmholtz instability broaden, but the frequencies corresponding to the given solutions ($\mathrm{Im}\, S$) also increase. Otherwise the qualitative relations describing the behavior of the $S = S(U)$ curves are similar to those shown in Fig. 2.

The $S = S(U)$ curves shown in Figs. 4a and 4b exhibit something of an exotic character in that they have been calculated for equal densities of the medium and the drop. Here not only is interaction observed between consecutive modes — the second and third (curve *8*) and the sixth and seventh (curve *11*), but branches corresponding to interaction of the second and fourth modes (curve *9*) and third and fifth modes (curve *10*) are also given. Nonetheless, such a dependence can be realized for physical entities such as ball lightning and help to explain why there have been no accounts of rapidly moving ball lightning in eyewitness reports.[14,15]

When a charge is present on a drop (Fig. 5), the main trends of the $S = S(U)$ curve are as described above except for a decrease in the critical velocities at which the drop becomes unstable, consistent with the analytical relation (19) and the data in Fig. 1. The curves are numbered the same as above.

Figure 6 illustrates the case of a drop carrying a charge slightly above the critical value for Rayleigh decay: $W = 4.5$ [the drop becomes unstable at $W = 4$ (Ref. 3)]. It is evident that an unstable solution characterized by the growth
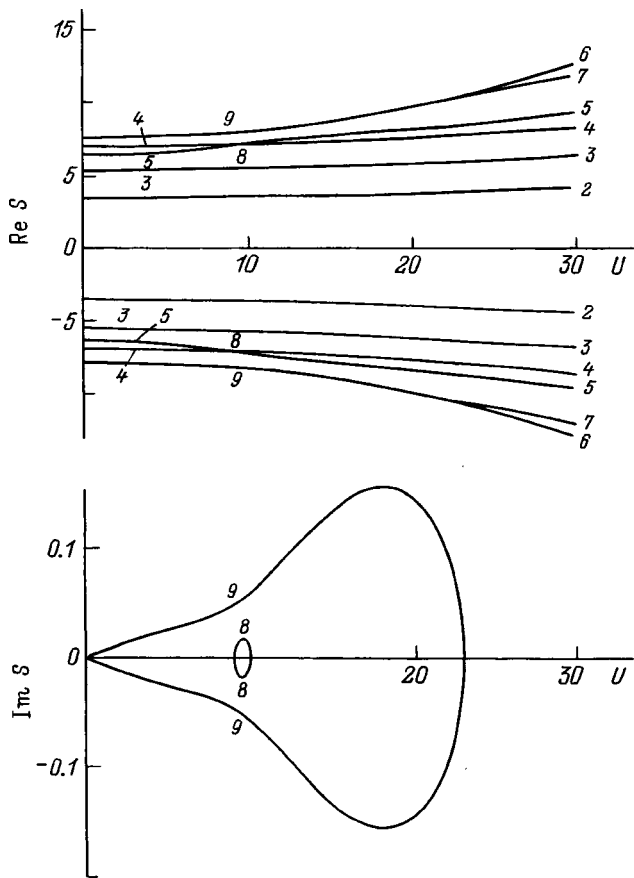
FIG. 7. Dependence of the real and imaginary components of the frequency S on the velocity U for $\rho = 10^{-3}$ and a highly Rayleigh-supercritical charge $W = 10$. Curves *8* and *9* are the result of mode interaction: *8* — between the fourth and fifth modes; *9* — between the sixth and seventh modes.

rate of the second mode exists at low passing flow velocities. Otherwise the given $S = S(U)$ is identical to the one shown in Fig. 5. The critical velocities at which the drops become unstable are further reduced in this case.

The relation $S = S(U)$ for a highly supercritical charge $W = 10$ is shown in Fig. 7. This situation differs from the one discussed above in that the fundamental mode has the lowest instability growth rate among the modes, and the instability of the higher modes, which, depending on the velocity U, is oscillatory or aperiodic in character, governs the breakup of the drop.

The emergence of a high electric charge on a drop indicates that two different types of instability can occur in it: instability with respect to self-charge[3,11] and instability with respect to the tangential discontinuity of the velocity field at the free surface.[8,9,13] The interaction of these instabilities creates several different channels for breakup of the unstable drop. In any case the daughter droplets now carry high charges regardless of the channel by which breakup occurs.

5. A charged drop immersed in a flow of an ideal liquid or gas can become unstable and emit highly charged droplets at values of the charge which are subcritical in the sense of instability with respect to self-charge (Rayleigh instability). This phenomenon is possible by virtue of the superposition of two types of instability: instability of the free surface of the drop with respect to the tangential discontinuity of the velocity field and instability with respect to the self-charge. The flow past the drop can acquire both aperiodic and oscillatory instabilities, depending on the density ratio of the drop and the charge, the charge of the drop, and the flow velocity past the drop. Aperiodic instability can be achieved by deformation to a prolate spheroid and breakup of the drop into two droplets of comparable size (at low passing flow velocities) or by deformation to a parachute configuration and breakup into a set of small and slightly larger droplets (at high passing flow velocities). The results of the foregoing analysis are in agreement with experimental data.[4]

[1] *Monodispersal of Matter: Principles and Applications*, edited by V. A. Grigor'ev [in Russian], Énergoatomizdat, Moscow, 1991.

[2] S. I. Shevchenko, A. I. Grigor'ev, and S. O. Shiryaeva, Nauchn. Priborostr. **1**, 3 (1991).

[3] A. I. Grigor'ev and S. O. Shiryaeva, Izv. Akad. Nauk, Mekh. Zhidk. Gaza No. 3, 3 (1994).

[4] A. L. Gonor and V. Ya. Rivkind, Itogi Nauki Tekh. Ser. Mekh. Zhidk. Gaza **17**, 98 (1982).

[5] V. M. Muchnik and Yu. S. Rud'ko, Tr. UkrNIGMI, No. 13, 96(1971).

[6] V. A. Dyachuk and V. M. Muchnik, Dokl. Akad. Nauk SSSR **248**, 60 (1979).

[7] A. I. Grigor'ev and S. O. Shiryaeva, Phys. Scr. **54**, 660 (1996).

[8] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics*, 2nd ed., [Pergamon Press, Oxford–New York (1987); Nauka, Moscow (1986), 439 pp.].

[9] S. O. Shiryaeva and A. M. Grigor'ev, *Methods for the Calculation of Electrodynamic Instability Conditions* [in Russian], Izd. Yaroslav. Gos. Univ., Yaroslavl, 1996.

[10] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum* (World Scientific, Singapore, 1988; Nauka, Leningrad, 1975).

[11] A. I. Grigor'ev and S. O. Shiryaeva, Zh. Tekh. Fiz. **61**(3), 19 (1991) [Sov. Phys. Tech. Phys. **36**, 258 (1991)].

[12] A. I. Grigor'ev and S. O. Shiryaeva, Zh. Tekh. Fiz. **59**(5), 6 (1989) [Sov. Phys. Tech. Phys. **34**, 502 (1989)].

[13] A. I. Grigor'ev and S. O. Shiryaeva, Zh. Tekh. Fiz. **66**(2), 23 (1996) [Tech. Phys. **41**, 124 (1996)].

[14] I. P. Stakhanov, *Physical Nature of Ball Lightning* [in Russian], Atomizdat, Moscow (1979), 242 pp.

[15] A. I. Grigor'ev, I. D. Grigor'eva, and S. O. Shiryaeva, in *Plasma Chemistry* [in Russian], Énergoatomizdat, Moscow (1991), pp. 218–248.

Translated by James S. Wood

Asymmetric interaction of blast waves and shock waves with a body flying at supersonic velocity

V. P. Goloviznin and I. V. Krasovskaya

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia

The problems of asymmetric interaction of a blunt wedge traveling at supersonic velocity with a cylindrical blast wave from a point explosion and with a plane shock wave are investigated by numerical simulation. The evolution of the interaction flow is analyzed, and data are obtained on how the structure of the shock layer changes. © *1999 American Institute of Physics.*
[S1063-7842(99)00305-0]

INTRODUCTION

The last few years have witnessed a steady interest in problems relating to the interaction of bodies moving at supersonic velocity with various types of disturbances in the incident flow,[1–12] specifically with an explosion zone.[1,3–7] One of the most frequently encountered gasdynamic models of explosion is the instantaneous point explosion model,[13] wherein a shock wave, bounding a gas volume in which the distribution of the parameters is markedly inhomogeneous in both space and time, propagates from the center into the surrounding space. The main mass of the gas trapped by the shock wave is concentrated near the wave front, where very large pressure and density gradients are observed. The central zone is characterized by an almost total absence of gas and by a very high temperature. Transient phenomena stimulated by the supersonic motion of a body through an explosion zone produce significant changes in the structure of the flow around the body and impose dynamical and thermal loads on the body. Axisymmetric interaction flows have been investigated in several papers.[1,3–6] The results of calculations of the encounter of a supersonic ($M_\infty = 2.95$), spherically blunted cylinder with a low-intensity ($M_S = 1.2$) plane blast wave are given in Ref. 1. The gas dynamic parameters behind the blast wave front were assumed to be constant in the calculations. In this case a new steady flow begins to evolve around the bow section of the body, corresponding to the flow parameters behind the incident wave front.

The results of calculations of the flow of interaction of a supersonic ($M_\infty = 2.0$) blunt cone with a with a spherical blast wave ($M_S = 4.37$) from a point explosion are given in Ref. 6. The flow around the body remains extremely unsteady throughout the entire interaction stage. As the cone moves deeper into the interior of the explosion zone, the forward part of the bow shock rushes away from the body, causing the bow shock to stretch out toward the center of the explosion. The pressure ratio at the symmetry point of the bow shock tends asymptotically to unity.

Turchak and Kamenetskiĭ[3,4] have numerically modeled the interaction of a supersonic, spherically blunted body with a spherical blast wave emanating from a point explosion and having, at the instant of encounter, a radius comparable with the blunting radius of the nose of the body. Calculations were carried out up to the time at which the body passes through the center of the explosion and overtakes the wave front of the undisturbed blast wave. It was shown that secondary rarefaction and shock waves are formed inside the shock layer in this case.

Data on asymmetric interaction are all but lacking, being represented solely by a calculation[1] of the lateral effect of a plane shock wave ($M_S = 1.1$) on a supersonic ($M_\infty = 2.95$), spherically blunted cylinder.

In this paper we discuss the results of numerical simulation of the asymmetric interaction of blast waves and shock waves with a blunt wedge moving at supersonic velocity. We consider the problem for the planar, two-dimensional case.

STATEMENT OF THE PROBLEM AND NUMERICAL METHOD

Figure 1 shows an interaction flow diagram at the time $t = t_0$ for the case when the oncoming wave is a blast wave. A cylindrically blunted, infinite wedge moves with supersonic velocity through a homogeneous medium, which is at rest. The velocity vector of the wedge is parallel to its symmetry axis. A steady flow moves past the wedge with an outgoing bow shock. At $t = t_0$ a point C of the bow shock, not a symmetry point, hits the wave front of a blast wave of radius $R(t_0)$ propagating from a cylindrical point explosion in the same medium at rest relative to the explosion center O'.

The problem is solved in Cartesian coordinates $\{x, y\}$, whose origin is located at the center of the cylindrical blunting of the wedge and whose x axis coincides with the symmetry axis of the wedge. The medium is an ideal gas with a constant specific heat ratio γ. To describe the unsteady, planar interaction flow, we use Euler equations of motion written in the form of integral equations expressing the conservation laws.[5] To form dimensionless variables, the pressure p and the density ρ are scaled to the pressure p_∞ and the density ρ_∞ in the undisturbed body and in the explosive atmosphere, the velocities are scaled to $(p_\infty/\rho_\infty)^{1/2}$, linear dimensions are scaled to $(E/\alpha_0 p_\infty)^{1/\nu}$, and time is scaled to $(p_\infty/\rho_\infty)^{-1/2}(E/\alpha_0 p_\infty)^{1/\nu}$. Here E is the linear energy density of the cylindrical point explosion, $\alpha_0 = \alpha_0(\nu, \gamma)$ is a self-similar constant,[13] and ν is the symmetry parameter ($\nu = 2$ for a cylindrical explosion).
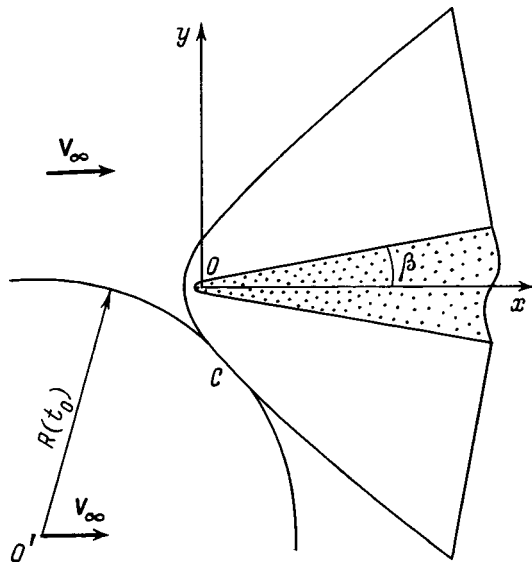
FIG. 1.

FIG. 2.

The system of equations of motion is solved numerically.[5] The bow shock is treated as a mathematical discontinuity (singled out), and all the other shocks are computed by the shock capturing technique. The calculations are carried out on a 68×20 grid.

The governing parameters of the problem are the Mach number M_∞ of the flow around the wedge before encounter, the Mach number M_B of the blast wave front at the instant of encounter, the specific heat ratio γ of the gas, the coordinates x_0, y_0 of the explosion center O', the half angle of the wedge β, and the ratio $\lambda = R(t_0)/r$ of the initial radius of the blast wave to a characteristic dimension (in this case the cylindrical blunting radius) of the wedge.

NUMERICAL RESULTS

Here we give the results of calculations for $M_\infty = 2$, $M_B = 4.841$, $x_0 = -0.06345$, $y_0 = -0.09807$, $\beta = 10°$, $\lambda = 40.13$, and $\gamma = 1.4$. At $t = t_0$ parameters corresponding to the solution of the problem of an instantaneous point (cylindrical) explosion propagating in a medium with counterpressure[13] are specified in the interior of the explosion zone. Figure 2 shows the pressure and density distributions along the radius of the explosion $R(t_0)$ at the time of encounter t_0. At the wave front we have $P_B = 22.62$ and $\rho_B = 4.777$. The nature of the pressure and density distributions behind the wave front of the incident blast wave corresponds to the earlier stage of the explosion with counterpressure, when the solution is qualitatively similar to the self-similar solution of the problem of a powerful explosion.[13] Sharp gradients of p and ρ occur in the vicinity of the wave front, where the main mass of the gas disturbed by the explosion is concentrated. The central zone of the explosion is characterized by a high temperature and a ''plateau'' in the distribution of the pressure along the space coordinate. The dynamics of the process of transient interaction of the supersonic body with the explosion zone is illustrated in Figs. 3 and 5, which show the calculated fields of equal densities at two
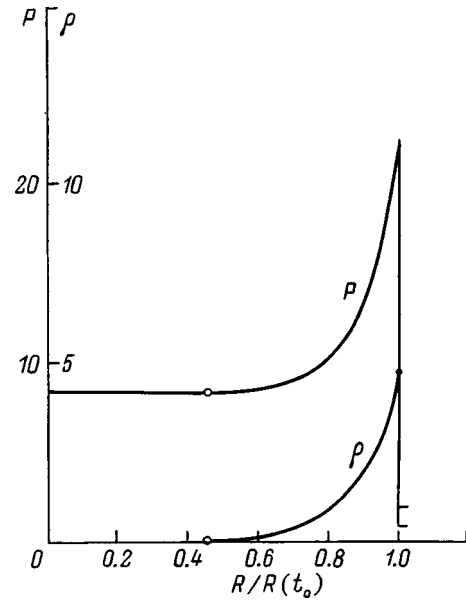
successive times. The dashed curve in each figure indicates the boundary of the central high-temperature zone of the explosion. The interaction flow diagram corresponding to Fig. 3 is shown in Fig. 4. At a time $t = t_1$ (Figs. 3 and 4) the wave front of the cylindrical shock wave, interacting with the bow shock, reaches the surface of the wedge and is reflected from it. For the lower part of the bow shock (below the axis of the wedge) the interaction with the blast wave front exhibits a regular behavior. The waves intersect at the point T_1 (Fig. 4), from which emanate two reflected shock waves and a contact discontinuity. An irregular interaction occurs when the incident blast wave runs into the top of the bow shock. A five-shock configuration is formed with two triple points T_2 and T_3, both of which are distinctly visible in Fig. 3.

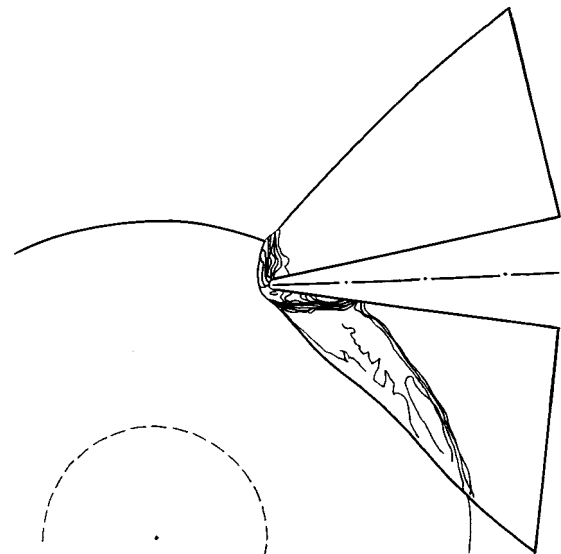Proceeding from the point T_1, the refracted part of the blast wave moves within the shock layer and is incident on
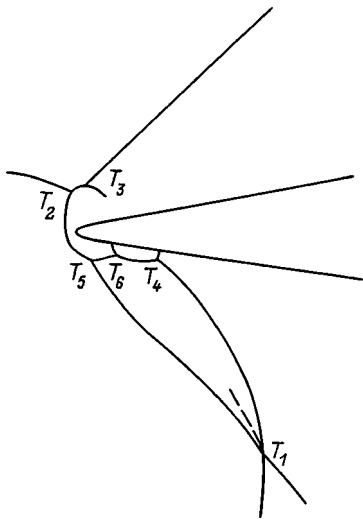

FIG. 3.

FIG. 4.

FIG. 6.

the lower surface of the wedge. A Mach-type reflection now takes place from the triple point T_4. As the refracted blast wave emanates from the point T_3, it has a low intensity and is not identifiable in Fig. 3. Simultaneously with the intersection of the shock waves, a new flow pattern begins to form around the nose section of the wedge, specifically an unsteady, divergent, floating flow behind the wave front of the incident cylindrical blast wave. The velocity of the gas at each point of this flow is the sum of the velocity of the gas relative to the explosion center O' and the translational velocity of the center relative to the xy coordinate system. It is evident in Fig. 3 that supersonic flow around the nose takes place at an angle of attack.

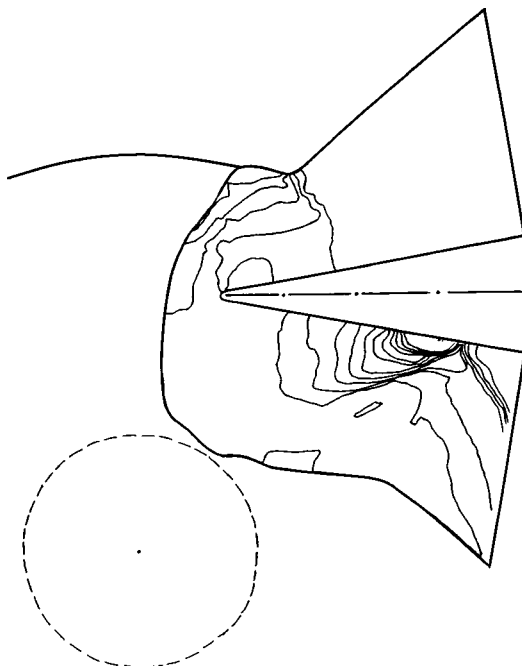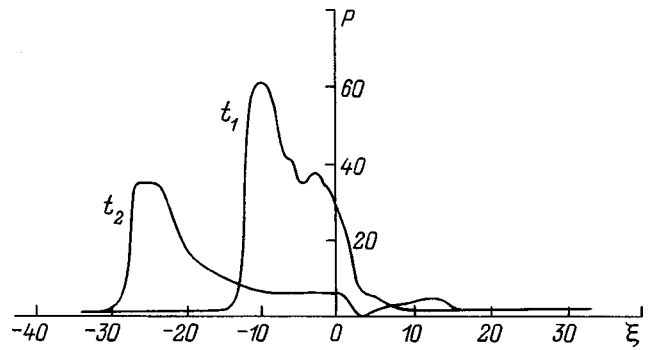Additional shock-wave configurations are formed as a result of interference of the two processes — flow around the body and the collision of the shock wave fronts. A triple point T_5 appears on the bow shock as a result of the irregular head-on collision of the shock waves, and another shock wave emanates from this point. The new shock wave, in turn, intersects with the wave reflected from the lower surface of the wedge, forming still another triple-shock configuration with the triple point T_6 (Figs. 3 and 4).

As the body moves through the interior of the explosion zone, the conditions of flow around the forward section of the wedge change, and the blast wave front weakens. An increase in the standoff distance of the bow shock from the body is noted at a time t_2 (Fig. 5). The lower part of the bow shock relative to the axis of the wedge moves through a gas of lower density and rapidly departs from the body, its intensity diminishing. The elongation of the bow shock in the direction of decreasing density, i.e., toward the explosion center O', is evident in Fig. 5. Simultaneously the shock structures with the triple points T_5 and T_6 disappear. The interaction of the incident blast wave with the upper part of the bow shock is still an irregular five-shock process and is accompanied by an increase in the length of the Mach stem.

Figure 6 shows a graph of the pressure distribution over the surface of the body at the two times t_1 and t_2. The coordinate

$$\xi = \frac{4}{\pi/2 - \beta} \frac{l}{r}$$

(l is the distance along the surface of the wedge) is measured from the forward critical point of the wedge. The positive direction of ξ corresponds to the upper face of the wedge. The maximum peak pressure occurs on the lower surface of the wedge after the Mach stem when the refracted blast wave is reflected from the surface of the wedge. The pressure distribution at time t_1 confirms the presence of a complex shock-wave flow structure about the nose of the wedge. The flow stagnation point shifts from the forward critical point of the body ($\xi = 0$) to the lower face of the wedge ($\xi \cong -3$).

At time t_2 the peak pressure after the Mach stem decreases. This effect is attributable to the decay of the blast wave intensity with time. The presence of an almost constant-pressure zone on the surface of the wedge is explained by the entry of the wedge into the explosion zone, where small disturbances propagate with very high velocities, rapidly equalizing the pressure.
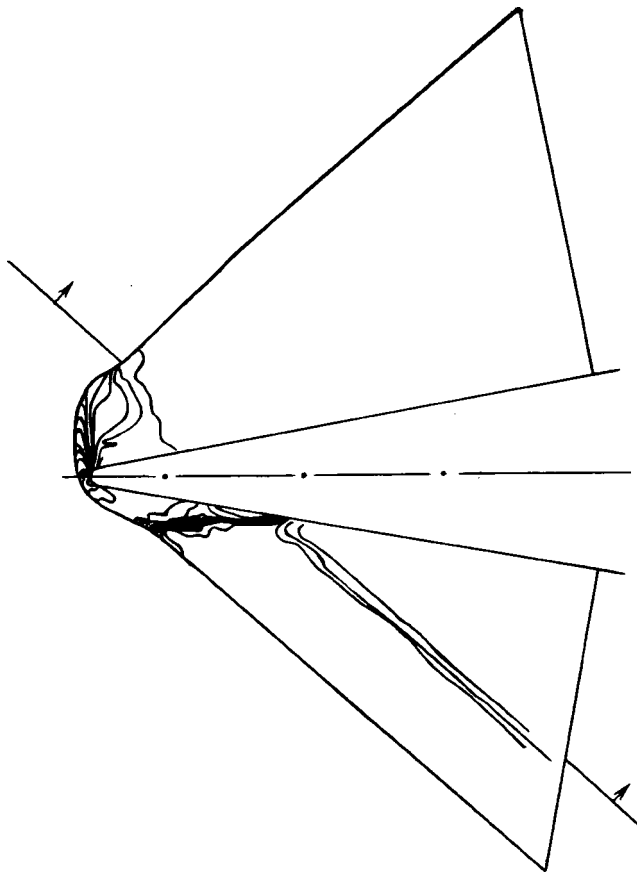

FIG. 5.

FIG. 7.

FIG. 8.

We have also investigated the asymmetric interaction problem for the case when the oncoming incident wave has a plane front and constant gasdynamic parameters behind the front. The governing parameters in this case are M_∞, M_B, β, γ, and the angle ω between the axis of the wedge and the shock wave front. Here we give the results of calculations for $M_\infty = 2$, $M_B = 4.37$, $\beta = 10°$, $\gamma = 1.4$, and $\omega = -43°$. For the given angle ω the wave front of the incident shock wave before encounter is parallel to the lower straight part of the bow shock. Figures 7 and 8 show the flow patterns in the form of equal-density lines at two times. The shock-wave interactions and the nature of the flow around the nose of the wedge at time t_1 are similar in this and in the preceding example (Figs. 3 and 7). The later time is characterized by the stable onset of steady flow around the wedge at an angle of attack, because the flow parameters behind the incident shock wave are constant (Fig. 8). We call attention to the configuration formed in the interaction of the incident shock wave with the upper part of the bow shock. The local interaction parameters (the intensity of the colliding waves and the angle between them) are such that the encounter diagram is necessarily irregular. A distinct five-shock configuration with two triple points is not evident in Fig. 8. The angle between the incident shock wave and the refracted part of the bow shock is 43° and, according to Landau's concept of shock directionality, this refracted shock is ''incoming'' for the triple point T. Most likely a compression pulse 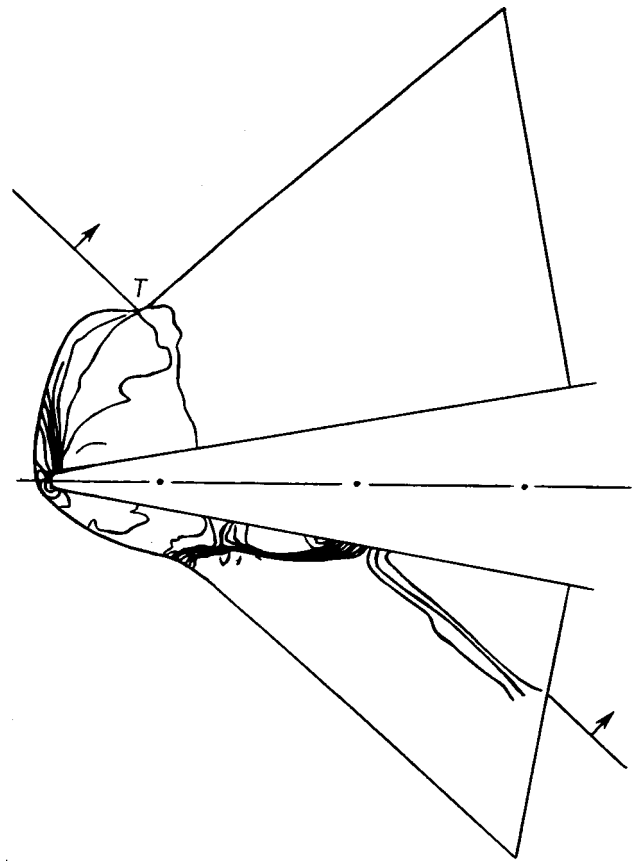transmitted from the vicinity of the nose of the wedge attains the triple point, inducing such a deformation of the shock-wave structure.

We close with the remark that the proposed method can be used to reliably simulate complex gasdynamic flows on coarse computational grids and also to investigate transient processes of interaction of supersonic bodies with various inhomogeneities.

[1] J. Champney, D. Chaussee, and P. Kutler, AIAA Paper 82-0027 (1982).

[2] V. F. Kamenetskiĭ and L. I. Turchak, Izv. Akad. Nauk SSSR Mekh. Zhidk. Gaza, No. 5, 141 (1984).

[3] V. F. Kamenetskiĭ and L. I. Turchak, in *Numerical Simulation in Aerohydrodynamics* [in Russian], Nauka, Moscow (1986), pp. 104–115.

[4] L. I. Turchak and V. F. Kamenetsky, in *Lecture Notes in Physics*, No. 323 (Springer-Verlag, Berlin-New York, 1989), pp. 578–585.

[5] V. P. Goloviznin and I. V. Krassovskaya, in *Gas Dynamics*, edited by Yu. I. Koptev (Nova Sci. Publ., 1992), pp. 191–229.

[6] V. P. Goloviznin and I. V. Krasovskaya, Zh. Tekh. Fiz. **61**(12), 12 (1991) [Sov. Phys. Tech. Phys. **36**, 1332 (1991)].

[7] V. P. Goloviznin and I. V. Krasovskaya, Zh. Tekh. Fiz. **64**(6), 38 (1994) [Sov. Phys. Tech. Phys. **39**, 540 (1994)].

[8] Yu. P. Golovachev and N. V. Leont'eva, Zh. Vychisl. Mat. Mat. Fiz. **28**, 148 (1989).

[9] Yu. P. Golovachev and N. V. Leont'eva, Zh. Vychisl. Mat. Mat. Fiz. **29**, 791 (1989).

[10] P. A. Voĭnovich, A. A. Fursenko, and S. V. Yuferev, FTIAN SSSR Preprint No. 1321 [in Russian] (Physicotechnical Institute, Academy of Sciences of the USSR, Leningrad, 1989).

[11] V. P. Goloviznin, G. I. Mishin, Yu. L. Serov, and I. P. Yavor, Zh. Tekh. Fiz. **57**, 1433 (1987) [Sov. Phys. Tech. Phys. **32**, 853 (1987)].

[12] V. U. Nabiev and S. V. Utyuzhnikov, Prikl. Mat. Mekh. **60**, 613 (1996).

[13] Ch. S. Kestenboĭm, G. S. Roslyakov, and L. A. Chudov, *The Point Explosion: Computational Methods and Tables* [in Russian], Nauka, Moscow (1974), p. 255.

Translated by James S. Wood

Effect of cathode surface state on the characteristics of pulsed hollow-cathode glow discharges

N. V. Gavrilov and S. E. Romanov

Institute of Electrophysics, Urals Branch of the Russian Academy of Sciences, 620049 Ekaterinburg, Russia
(Submitted October 6, 1997)
Zh. Tekh. Fiz. **69**, 20–24 (May 1999)

A study is made of the effect of pulse repetition rate ($0.1–10^3$ s^{-1}) and average discharge current (0–1 A) on the breakdown delay time and burning voltage of low-pressure glow discharges ($p < 0.1$ Pa) in an electrode system of the reverse magnetron type with a large cathode surface area ($\approx 10^3$ cm^2). It is shown that increasing the repetition rate leads to a many-fold reduction in the statistical spread in the delay time and in the discharge formation time, while the average discharge current has a significant effect on the burning voltage. The mechanism for the observed phenomena is interpreted qualitatively in terms of the presence of thin dielectric films on the cathode surface. © *1999 American Institute of Physics.* [S1063-7842(99)00405-5]

INTRODUCTION

Low-pressure reflex discharges inside a hollow cathode in a low magnetic field are operated in a repetitive-pulsed mode in plasma sources for large electron and gas ion beams.[1,2] The discharge is excited in an electrode system with a large interelectrode gap (>10 cm) at low gas pressures $p < 0.1$ Pa by applying a relatively low pulsed voltage U of amplitude 2–3 kV. The large area of the working surface of the hollow cathode ($\sim 10^3$ cm^2) and the low duty cycle of the pulses mean that the average ion current at the cathode is low, although the pulsed current density is usually 1–10 mA/cm^2 for beam current pulse durations $\tau \sim 0.1–1$ ms and repetition rates $f \sim 1–100$ s^{-1}. Conventional oil vacuum pumps can be used to obtain the vacuum in sources for technological purposes. Because the samples have to be changed periodically, the electrodes in the source come into frequent contact with the atmosphere. Thus, despite partial glow-discharge conditioning of the cathode surface, both adsorbed gas atoms and contaminants may be on it.

Under these conditions the experimentally measured delay times to breakdown are rather long, from tens to hundreds of microseconds. Experiments[3,4] have shown that the delay time depends on the applied voltage, gas pressure, and magnetic induction, which determine the rate of rise of the current in the gap during the development of Townsend avalanches, but also on such parameters of repetitive-pulsed operation as the pulse repetition rate f and the current pulse duration τ and amplitude I. A dependence of the measured delay time on the prehistory of the discharge and the repetitive-pulsed operating parameters for other discharge burning conditions has been observed before.[5,6] However, in all the previous reports, changes in the statistical average values of the delay time were studied, while in our experiments[4] a substantial change in the delay time was observed with a minimal statistical scatter in the measured values. It was shown[4] that the burning voltage of pulsed discharges also depend on the pulse repetition rate, as well as on

the magnetic induction B, pressure p, and current amplitude I. Since changes in the ignition and burning conditions for pulsed discharges have a significant effect on the operation and parameters of beams from charged-particle sources based on these discharges, studies of the discharge characteristics in the repetitive-pulsed mode are essential.

In this paper we study the effect of the pulse repetition rate and average discharge current on the delay time and burning voltage of a large hollow cathode glow discharge and propose a qualitative explanation for processes on the glow discharge cathode which are capable of affecting the properties of the pulsed discharges under experimental conditions.

EXPERIMENT

In the experiments we used the electrode system of an ion source[1] (Fig. 1) consisting of a cylindrical hollow cathode *1* of stainless steel with equal length L and diameter D (150 mm) and a rod anode *2* of tungsten with a diameter $d = 3$ mm and length $l = 100$ mm located on the system axis. Gas (argon, nitrogen, oxygen) was fed directly into the cathode hollow and pumped out through the end *3* of the cathode, which was covered with a fine (0.6×6 mm) stainless-steel grid. The pressure of the gas in the vacuum vessel for a gas feed of 1 cm^{3}atm/s into the discharge gap was 0.01 Pa. The vacuum was created by a oil-diffusion pump without a trap. A longitudinal magnetic field with an induction of 1 mT was created by solenoids *4*. A cylindrical Langmuir probe *5* with a diameter of 0.4 mm and length of 6 mm was mounted on the axis of the discharge system.

Discharges were excited in a repetitive-pulsed mode with a repetition rate that could be adjusted over the interval 0.1–1300 Hz. The pulse duration was varied over 0.1–1 ms and the pulse rise time was less than 5 μs. The pulse voltage amplitude U applied to the gap was varied over 1–3 kV. The instability in the voltage pulse amplitude and repetition rate was less than 5%. The discharge current was varied over

FIG. 1. A sketch of the electrode system.
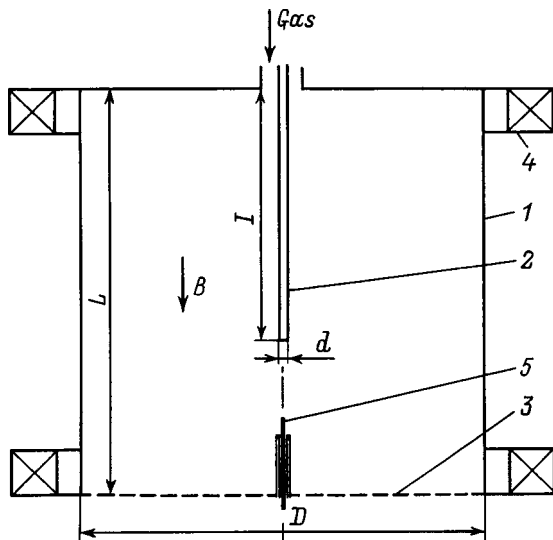

FIG. 3. Breakdown delay time as a function of discharge current pulse repetition rate: $U=2$ kV, $I=10$ A, $\tau=1$ ms.

0.05–20 A. In the experiments on the effect of the average discharge current on the burning voltage of the pulsed discharge, we used a cw discharge with a regulated current of 0.1–1 A. In this case a pulsed voltage was applied the gap with a dc discharge burning in it.

The breakdown delay time t was defined as the interval from the time the voltage U was applied until the time of the sharp drop in the voltage across the discharge gap to a level of $0.9U$. The burning voltage of the discharge was measured after the discharge current I reached its steady level. The data presented here were averaged over 64 oscilloscope measurements.

Before the measurements were made, the discharge system was aged for an extended period (~ 1 h) with a fixed value of the average discharge current until the parameters of the pulsed discharge ceased to vary. Figure 2 shows plots of the breakdown delay time t as a function of the time $1/f$ between discharge current pulses in argon after aging at average currents of 0.5 A (curve 1) and 0.1 A (curve 2). An
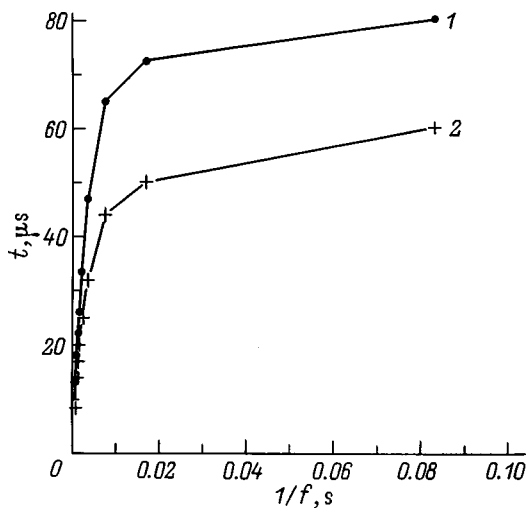
increase in f leads to a sharp reduction in the delay time t, and t also depends on the preliminary aging conditions. At a frequency of ~ 5 s^{-1} the statistical scatter in the values of the delay time has fallen to a minimum ($\sim 2\%$) owing to the instability in the amplitude and repetition rate of the pulses and in the discharge formation time (Fig. 3). Besides a reduction in the statistical spread over frequencies $f<5$ s^{-1}, a decrease in the minimum detected delay times was also observed. As the frequency is raised further, the delay time continues to decrease monotonically and the scatter in t remains minimal. After replacement of argon by nitrogen or oxygen and extended aging of the cathode, the $t(f)$ curves and the variations in the scatter in t with frequency are essentially the same as for the argon discharges. The distinctive feature of repetitive-pulsed discharges in oxygen is a much lower rate of increase in t for $f<200$ s^{-1}.

The current–voltage characteristic of discharges in inverted magnetron electrode systems is rising and the differential resistance dU/dI of the discharge decreases with rising current. Thus, in the pulsed high current regime ($I\sim 5$–10 A), the discharge burning voltage depends weakly on the current amplitude, but its magnitude can be either higher or lower than the burning voltage of dc discharges with lower currents for the same conditions in the gap. The magnitude of the average discharge current has a more important effect on the discharge burning voltage in the repetitive-pulsed mode. Figure 4 shows some data illustrating the effect of the average current in the discharge gap on the burning voltage of pulsed and dc discharges. A pulsed voltage was applied to the gap with a dc discharge in it at a current of 0.1–1 A and a discharge with a frequency $f=10$ s^{-1} and a current amplitude of 10 A was excited. As the dc current was raised, the voltage of both the dc and the pulsed discharges increased. After the dc discharge was extinguished, the burning voltage of the pulsed discharge fell to the lower initial values after a few minutes. Over the first

FIG. 2. Breakdown delay time as a function of the time between discharge current pulses: $U=2$ kV, $I=2$ A, $\tau=0.1$ ms.
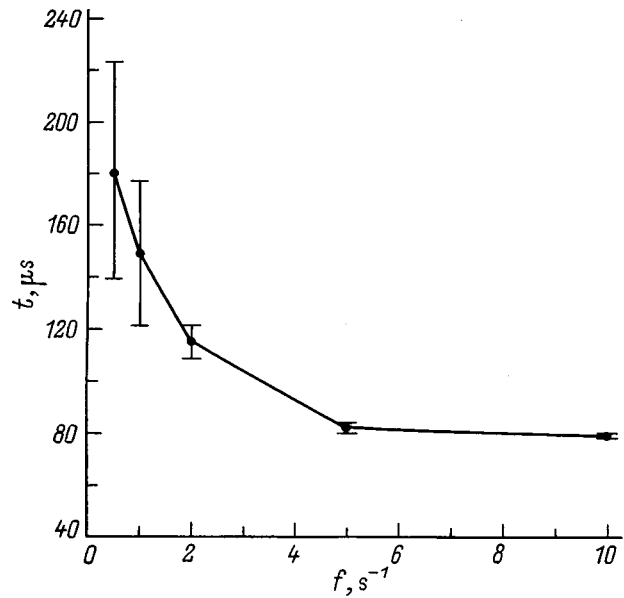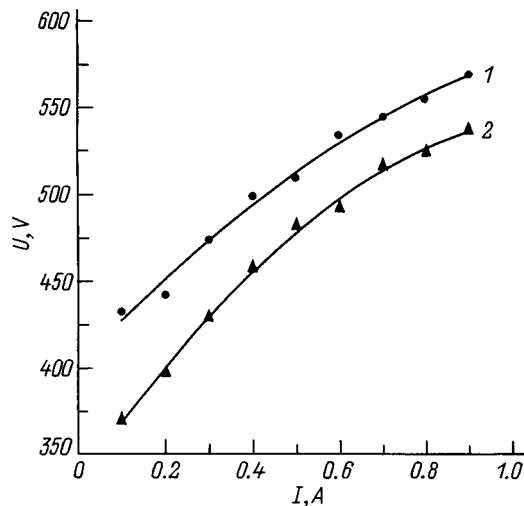
FIG. 4. Burning voltages of pulsed (*1*) and dc (*2*) discharges as functions of the dc discharge current. The parameters of the repetitive-pulsed discharge are $I = 10$ A, $f = 10$ s^{-1}, and $\tau = 1$ ms.

1–2 min the discharge delay time rose slightly (by ~ 5 μs). Increasing the current pulse repetition rate leads to a rise in the burning voltages of both discharges.

The time constant for decay of the plasma was determined from the rate of change, after the discharge was turned off, of the ion current in the circuit of a Langmuir probe that was negatively biased relative to the cathode. The time for the voltage across the gap to fall to zero was a few μs. The estimated decay time for the discharge lies within 15–20 μs.

DISCUSSION OF RESULTS

The discharge delay time t is determined both by the rate ν at which initiating electrons appear in the discharge gap and by the formation time of a glow discharge structure in the gap as the result of the development of a series of Townsend avalanches. A decrease in the statistical spread in t with increasing repetition rate (Fig. 3) is evidence of a higher current of initiating electrons. In the experiments, this current may be caused by the presence of residual charges in the gap, the influence of metastable atoms and molecules, or post-discharge electron emission from the cathode.[7] Since the drop in the plasma density in the volume as the plasma decays is caused by wall recombination, the density fall must be exponential, as confirmed by experiment. In this case, estimates show that the effect of residual charges can show up at repetition rates on the order of 1 kHz for a constant plasma decay time of ~ 30 μs. Since a change in the delay time is observed at much lower repetition rates, while the measured plasma decay times are 15–20 μs, we conclude that the behavior of the $t(f)$ curves is not caused by residual charges in the discharge gap. It is also improbable that metastables exist on the cathode surface for such long times (~ 0.1 s).

Under the experimental conditions the most probable reason for the appearance of initiating electrons in the discharge gap may be post-discharge electron emission from the cathode. The characteristic drop in the current with time for post-discharge emission can explain the behavior of the

$t(f, 1/f)$ curves shown in Figs. 2 and 3. The reduction in the rate ν at which initiating electrons appear as the pause between current pulses is increased leads to a rise in the average values of t and to their statistical spread. The present experiments are distinguished by the use of a cathode with a large working surface area, which has led to a reduction in the statistical spread in the values of the delay time and has permitted direct measurements of the discharge formation time at the low current densities following emission ($10^{-17} - 10^{-16}$ A/cm^2) from a cathode surface that has been partially conditioned in the discharge and made it possible to study the dependence of the discharge formation time on the conditions in the gap and on the parameters of the repetitive-pulsed mode.

The statistical spread in t falls to a minimum for $f > 5$ s^{-1} (Fig. 3). The further reduction in t with increasing f may be caused only by a change in the formation time of a glow discharge. As the rate f is changed from 10 to 1300 Hz, the discharge formation time in argon falls several fold (from 60–80 to 10–15 μs under the conditions of Fig. 2). The average discharge current during preliminary aging of the cathode has the opposite effect on t, i.e., increasing the current leads to a drop in t and reducing the current, to a drop in the delay time, but the changes in t in this case are less, 5–15 μs, and take place over a rather extended time (several minutes). This dependence of t on the average current indicates that aging is not equivalent to an ordinary cleaning of contamination from the cathode surface, but provides for the establishment of some equilibrium state of the surface after a certain relaxation time.

An increase in the burning voltage of the pulsed discharge is observed when the dc discharge current is increased within the same interval (Fig. 4), as well as when the average current of the repetitive-pulsed discharge is increased by varying the pulse repetition rate.[4] Turning off the dc discharge leads to recovery of the initial voltage and delay time after a few minutes. After the steady state of the cathode surface has been attained, no significant change in the parameters of the repetitive-pulsed mode is observed over an extended period (several hours).

These data imply that changing the average discharge current has a significant effect on the state of the cathode surface, in which both the ion–electron emission and the electron after emission coefficient of the cathode change. Here the means by which the magnitude of the average current is changed (by varying the pulse repetition rate or by varying I, τ, and dc discharge current) has different effects on the variation in t and the same effect on the burning voltage of the discharge. This variation in the emission properties of the cathode can be explained by assuming that there are nonconducting thin films on the cathode surface, whose regeneration and properties are determined by the magnitude of the average discharge current.

It is known[8] that the effect of an ion flux on a metal surface depends on the relationship between the ion current density and the flux of residual gas particles adsorbed on the cathode surface. Under ordinary conditions, a regime of predominant sputtering and cleaning of the cathode is achieved only at rather high ion current densities of 0.1–1 mA/cm^2.

When chemically active impurities are present in the volume and the ion current density is low, a dynamic equilibrium may develop between the adsorption and sputtering processes such that a layer with a modified chemical composition and structure may exist on the cathode surface for a long time. Direct measurements of the composition of the cathode surface in a dc glow discharge have shown[9] that the reduction in the discharge current with time coincides with a lowering of the impurity content on the cathode surface, where the thickness of the contaminating oxygen- and carbon-containing film is estimated to be only a few tens of Å. Thin dielectric films and post-discharge emission at current densities of 10^{-18}–10^{-15} A/cm^2 have been observed in experiments[10] after all kinds of cleaning, including various combinations of high temperature annealing in vacuum and processing with glow discharge electrons.

The presence of nonconducting films on the cathode surface makes it possible to explain both the features of the ignition process and the dependence of the discharge voltage in the quasistationary phase of the discharge on the parameters of the repetitive-pulsed mode. The drop in the formation time t with rising pulse repetition rate f and the corresponding rise in the post-discharge emission current occur because when $t \gg 1/\nu$, many-electron ignition takes place. If the current in the gap rises in an avalanche, i.e., $I(t) = I(0)\exp[-t(\mu-1)/\tau]$, where τ is the time of flight of an ion and μ is the current gain in the gap, then a several-fold drop in the formation time means that the initial current $I(0)$ has increased by roughly the same order of magnitude. Thus, the experimentally observed drop in t means that the after-emission current density rises by 4–5 orders of magnitude while the time interval is reduced from 10^{-1} to 10^{-3} s.

The dynamics of the post-emission process has been studied by measuring the discharge delay times under various conditions and after various kinds of processing[5,7,10] over times ranging from a few seconds to tens of hours. All these data indicate a rise in the rate of change of the post-discharge emission current as the time from discharge cutoff is reduced. The high repetition rates used in our experiments allowed us to make measurements with a short time delay (10^{-3} s) relative to the time the discharge is quenched and to obtain a high rate of change of the post-discharge electron emission current for $1/f < 10^{-2}$ s.

It is known that adsorption of atoms of the principal contaminant impurities, oxygen and carbon, on a metal surface raises the work function of the metal and, accordingly, reduces the ion–electron emission coefficient γ.[11] Increased field emission of electrons can take place as a result of the formation of a thin insulating film on the metal surface that is permeable for ions bombarding the cathode and the formation, in that film, of an electric field which facilitates the escape of secondary electrons from the metal owing to the Schottky effect. This kind of field is created by the development of a positive electrical charge on the film surface owing to emission from the discharge and the deposition of positive ions.[5,7] Microscopic inhomogeneities on the cathode surface which aid in enhancing the field within the film cause post-discharge field emission from the cathode.

The rapid drop in the post-discharge emission current to

low levels indicates that there is little contamination of the cathode, which, nevertheless, has a significant effect on the properties of the pulsed discharge. It should be noted that the presence of a cathode surface layer with a variable composition and structure does not cause significant contamination of the beam of gas ions.[1]

CONCLUSION

The state of the cathode surface in a glow discharge under imperfect vacuum conditions depends on the average discharge current and has a significant effect on the ignition and burning parameters of a pulsed discharge. At low average discharge currents thin insulating films form on the cathode surface and they cause a significant reductions in the discharge delay time and the discharge burning voltage. As the average current is raised, the influence of the films is reduced. The relaxation time for the discharge parameters is a few minutes when the average current is varied.

A technique based on using a cathode with a large working surface area and varying the pulse repetition rate over wide limits has made it possible to study the dynamics of the variation in the post-discharge emission with high time resolution and to observe a sharp rise in the after emission current density when the time interval following quenching of the discharge is reduced. A high post-discharge emission current density from the large cathode surface at elevated current pulse repetition rates ensures many-electron initiation of the discharge, one consequence of which is a reduction in the discharge formation time.

The high degree of influence of thin films on the emission properties of the cathode of a glow discharge observed in these experiments makes it necessary to take the existence of these films into account in experiments with pulsed glow discharges in devices operating under commercial vacuum conditions.

[1] N. V. Gavrilov, G. A. Mesyats, S. P. Nikulin *et al.*, J. Vac. Sci. Technol. A **14**, 1050 (1996).

[2] N. V. Gavrilov, M. A. Zav'yalov, S. P. Nikulin *et al.*, Pis'ma Zh. Tekh. Fiz. **19**(21), 57 (1993) [Tech. Phys. Lett. **19**(11), 689 (1993)].

[3] N. V. Gavrilov and S. P. Nikulin, in *Proceedings of the XVI International Symposium on Discharges and Electrical Insulation in Vacuum*, edited by G. Mesyats, Proc. Soc. Photo-Opt. Intstrum. Eng. (SPIE) **2259** (1994), pp. 118–121.

[4] N. V. Gavrilov and G. Mesyats, in *Proceedings of the XXIII International Conference on Phenomena in Ionized Gases*, edited by M. C. Bordage and A. Gleizes, Univ. P. Sabatier, Toulouse (1997), Vol. 2, pp. 32–33.

[5] H. Paetow, Z. Physik, **111**, 770–791 (1939).

[6] V. Kudrle, E. Le Due, and M. Fitaire, in *Proceedings of the XXIII International Conference on Phenomena in Ionized Gases*, edited by M. C. Bordage and A. Gleizes, Univ. P. Sabatier, Toulouse (1997), Vol. 4, pp. 28–29.

[7] Yu. A. Korolev and G. A. Mesyats, *The Physics of Pulsed Breakdown in Gases* [in Russian], Nauka, Moscow (1991).

[8] N. V. Pleshivtsev, *Cathode Sputtering*, Atomizdat, Moscow (1968).

[9] J. E. Houston and R. D. Bland, J. Appl. Phys. **44**, 2504 (1973).

[10] P. N. Chistyakov and N. V. Tatarinova, Zh. Tekh. Fiz. **35**, 1333 (1965) [Sov. Phys. Tech. Phys. **10**, 1035 (1965)].

[11] M. Kaminsky, *Atomic and Ionic Impact Phenomena on Metal Surfaces* [Springer-Verlag, Berlin (1965); Mir, Moscow (1967), 506 pp.].

Translated by D. H. McNeill

Dynamic model for plasma opening switches

P. I. Zubkov

M. A. Lavrentev Institute of Hydrodyamics, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

A dynamic model for plasma opening switches is proposed. The basis of the model is the appearance and development of force instabilities (pinch and sausage) in a spatially nonuniform plasma accelerated by a magnetic field as a driver piston. The proposed model does not require the invocation of subtle effects in the pinch collapse phase and makes it possible, without going beyond the framework of magnetohydrodynamics, to explain the principal operating features of plasma opening switches and to obtain quantitative estimates consistent with experiment. © *1999 American Institute of Physics.* [S1063-7842(99)00505-X]

INTRODUCTION

Plasma opening switches are one of the most powerful generators of electrical pulses. Further increases in their power, matching them to loads, extending their range of applications, controlling their operation, etc., require the creation of a model that describes their operation in a satisfactory manner. One often cited model,[1] based on the formation of double layers and a carrier deficit, does not describe the main operating characteristics of plasma opening switches and has been shown[2–4] to be unsatisfactory. Other models[3–6] can explain only particular aspects of the phenomenon, without providing a complete picture of the processes in a plasma opening switch. They do not take the plasma dynamics and the mutual effect of the storage device and load on the plasma into account.

In this paper a dynamic model for plasma opening switches is proposed based on the appearance and development of force instabilities (pinch and sausage) in a spatially nonuniform plasma accelerated by a magnetic field. When the plasma in the pinches and sausages collapses, high voltage pulses are generated which switch the current from the storage device of the plasma opening switch into the load.

ESTIMATES OF CHARACTERISTIC PARAMETERS

As a rule, a plasma opening switch consists of a segment of coaxial line, one end of which is attached to an energy source or transmission line while the other is connected to a load. The load is usually a vacuumdiode and a capacitor bank with a capacitance C is used as an energy source. Plasma guns are positioned azimuthally in the outer electrode of the coaxial line and these inject plasma radially toward the central electrode. The inductance of the source and transmission line up to the plasma gun forms the inductive storage system L_0 of the plasma opening switch. The relative gap between the outer electrode (diameter D) and the inner diameter d, given by $(D-d)/2d$, lies between 0.5 and 10 or more for various designs. This leads to a difference of an order of magnitude or more between the magnetic pressures at the inner and outer electrodes. Estimates of the characteristic parameters are given here for an inner electrode radius $r \approx 2$ cm.

The major studies of plasma opening switches have been done for maximum currents of 10^4–10^6 A with a discharge half period of 2×10^{-7}–1.5×10^{-6} s and a current rate of rise of 10^{12}–10^{14} A/s. The operating period of the plasma guns is generally much longer than the time for the inductive storage unit to build up with the current. Despite the large number of papers on plasma opening switches, the injected plasma (mainly of carbon) has not been studied adequately. The input particle densities n are most often 10^{12}–10^{14} cm^{-3} (Refs. 1 and 7). In many papers the particle density is estimated using the model of Ref. 1. A number of papers[8] refer to an optimum density $\sim 10^{13}$ cm^{-3} at which a plasma opening switch still operates.

The velocity of the injected plasma is $\sim 5 \times 10^6$ cm/s.[1,9,10] If we assume that the plasma expands into a vacuum, then a good approximation for the sound speed in the absence of a magnetic field will be 10^6 cm/s, which yields an estimate of $T_0 \approx 5 \times 10^4$ K for the temperature. Here we assume the electron and ion temperatures are equal. In Ref. 7 the calculated temperature lies within $4 < T < 8$ eV. It is natural to assume that the injected plasma expands along the electrodes, as well. Then, by the time the inductive storage device is being fed by the current, its characteristic axial size will be $R \sim V\tau_0 \approx 10$ cm, where τ_0 is the delay between the time of plasma injection and the onset of filling. In some designs for plasma opening switches the expansion along the electrodes is limited by the design.

For the chosen plasma parameters the Debye radius is $\approx 10^{-4}$–10^{-3} cm, which is substantially smaller than any dimensions of the plasma, so it can be regarded as quasineutral. The electron mean freepath ~ 0.1–10 cm and may be on the order of the largest dimension of the plasma. The electron Larmor radius for a current rate of rise $\sim 10^{12}$ A/s is much shorter than the mean free path and any dimensions of the plasma $\sim 10^{-10}$ s after the current feed is turned on. This time is much shorter than the characteristic current rise times.

These estimates show that the possibility of a magneto-dynamic description of the plasma motion must be examined in detail for each specific case. As for the motion perpendicular to the magnetic field, because of the smallness of the Larmor radius, we shall assume that this approximation is satisfied.[11] In the following, essentially only this kind of motion is considered.

We shall assume that the plasma is fully ionized and its electrical conductivity is $\cong 10^7 \cdot T_0^{3/2} \approx 10^{14}$ s^{-1} (Ref. 11). During the current rise time of $10^{-7}-10^{-6}$ s, the magnetic field penetrates the motionless plasma to a depth of $\sim 0.3-1$ cm as a result of the skin effect, a distance much smaller than the size of the plasma along the storage device. Thus, we shall assume that the magnetic field has not penetrated the plasma.

The magnetic pressure $B^2/8\pi$ is comparable to the gas kinetic pressure ($nkT_0 \sim 10^2$ dyn/cm) for a rate of current rise of 10^{12} A/s over a time $\sim 10^{-9}$ s, which is at least two orders of magnitude shorter than the current buildup time in the storage system. Over a time of the same order of magnitude, a shock wave develops and moves through the unperturbed, current-free plasma. Thus, we shall assume that the current is distributed over a thin layer of compressed plasma and for purposes of estimates and a qualitative discussion, we use a ''snowplow'' model. The propagation of the magnetic field along the plasma opening switch will be determined by the plasma motion and not by diffusion.

The characteristic velocity according to a snowplow model with a sinusoidal current rise and a uniform distribution of the density ρ up to the time it reaches its maximum will be[12] $U = 2 \times 10^{-2} I_0/(\rho^{1/2}d)$, where I_0 is the current amplitude in kA. For the Gambl I system[13] with $n = 10^{12}$ cm^{-3}, this formula gives $U \approx 2 \times 10^8$ cm/s, or several orders of magnitude greater than the spin velocity, which eliminates an apparent contradiction between the experimentally measured propagation velocity of the magnetic field[13] (measured at $\approx 5 \times 10^8$ cm/s) and that calculated from its diffusion.

With these plasma parameters the mean free path for photons of all types is much greater than the plasma size, so we shall assume it to be transparent.

QUALITATIVE MODEL OF SWITCHING

The boundary of a uniform plasma accelerated by a uniform magnetic field parallel to it, is unstable. The acceleration dynamics is such that it initiates the appearance and development of the Rayleigh–Taylor instability.

In a plasma opening switch the plasma injected by the discrete plasma guns is spatially inhomogeneous. The distribution of the plasma density along the storage unit has been noted.[14] Because it is reflected from the inner electrode and because of the spreading geometry, the plasma will be inhomogeneous along the direction of injection, as well. It is natural to assume that it is also inhomogeneous along the azimuth, and this has been confirmed in direct measurements of the electron density.[7]

The inhomogeneity of the plasma density and the difference in the directions of its spread velocities lead to nonuniform acceleration of the front of the shock compressed plasma, to nonuniformities in the velocities gained by it, and to a complicated geometry of the plasma surface. The picture can be made worse by the order of magnitude difference between the magnetic pressures at the central and outer electrodes. Between the plasma guns, where the plasma density is lower, ''tongues'' of magnetic field are formed in the axial direction. After passing the planes of the plasma guns, the ''tongues'' of magnetic field that did not penetrate into the plasma begin to expand owing to the density nonuniformity and gradient acceleration and the lines of force can close and pinches can form.[15,16] It is well known[17] that a plane layer is unstable to pinches and sausages, even in the case of a homogeneous plasma. Pinch formation can be facilitated by the transparency of the plasma, as well. A transparent plasma is subject to the overheating instability, which develops into pinches.[17] In other words, we have all the conditions for the appearance and development of force instabilities and, in particular, pinches and sausages.

In the proposed dynamic model we assume that the plasma in the plasma opening switch is subject to pinching. We believe that the plasma is accelerated by the magnetic field along the storage unit and is compressed in pinches, with respect to which it is unstable. It is evident that this assumption requires more detailed examination and justification. However, assuming that pinches and sausages appear and develop makes it possible to explain qualitatively the main operational features of plasma opening switches and obtain some quantitative estimates of the characteristic parameters that are consistent with experiment.

The generation of charged particle beams may serve as an indirect confirmation of the appearance of pinches and sausages in the plasmas of plasma opening switches.[18] These beams were observed in the first investigations of Z-pinches in controlled thermonuclear fusion research.[19] The localization of the high electric fields observed experimentally in plasma opening switch plasmas[20] may be a consequence of the development of sausages and a confirmation of their possible appearance.

The collapse of the plasma in pinches leads to a rise in the induction of the conducting plasma channels. High voltage pulses, which switch the current from the storage system of the plasma opening switch to the load, develop across the increasing inductance. The expansion of the plasma begins to slow down and the pinches begin to be compressed when the condition

$$B^2/8\pi \geqslant p + \rho V^2 \tag{1}$$

is satisfied, where B is the magnetic induction at the pinch boundary and p, ρ, and V are the pressure, density, and expansion velocity of the plasma.

Expressing the induction in terms of the current I, we obtain

$$I \geqslant Rc(2\pi(p + \rho V))^{1/2}, \tag{2}$$

where R is the characteristic scale length of the plasma nonuniformity when a pinch develops and c is the speed of light.

Equations (1) and (2) are basic to a quantitative explanation of the mechanism for operation of plasma opening switches. Equation (2) implies that the proposed model is

indifferent to the direction of the flowing current and, therefore, to the polarity of the electrodes. It has been noted[21] that plasma opening switches operate with arbitrary combinations of the electrode polarities and plasma injection. The polarity effect[9] may be caused by an asymmetry in the motion of the compressed plasma front owing to the large difference in the magnetic pressures at the inner and outer electrodes, by the plasma gun current flowing through the plasma, and by continued injection of plasma into a region occupied by the magnetic field. The model of Ref. 1 allows a plasma opening switch to operate only with a single electrode polarity, where the outer coaxial electrode is the anode and the inner, the cathode, while the plasma is injected from the outer electrode.

Equation (2) explains the existence of the critical current I^*, known as the cutoff current. The cutoff current and its dependence on various factors have been studied by many authors. Equation (2) implies that for a small number K of plasma guns such that the plasmas from the individual guns do not overlap in a large part of the space between the electrodes, $I^* \propto K$. A dependence of this sort has been observed experimentally.[22] For a large number of plasma guns, the plasma flows overlap, so the inhomogeneity of the plasma is reduced and it becomes more stable. In this case, I^* increases and, therefore, the time before the onset of switching, or the plasma opening switch ceases to operate, entering a short-circuit regime.[8]

According to the model of Ref. 1, the critical current is determined by the ion velocity and the area covered by plasma on the central electrode but is independent of the number of plasma guns. Equation (2) implies that the cutoff current $I^* \propto n^{1/2}$. In many papers in the course of studying the dependence of the cutoff current on the plasma density it has been noted that it increases with rising plasma density and when the plasma opening switch operates longer. A comparison with experimental data is difficult because of the narrow range of the studies and of the unreliability of some of them.

Usually the energy source for driving the plasma opening switch is a capacitor bank which charges the inductance of the storage system. For more complete utilization of the stored energy the operating conditions are chosen so that the critical current I^* is essentially equal to the maximum current $I_0 \approx V_0 (C/L_0)^{1/2}$ in the ''hard'' circuit, where V_0 is the initial voltage on the capacitor. In this case $I_0 \propto n^{1/2}$. Many researchers have noted that raising the plasma density above the optimum leads to cessation of operation of the plasma opening switch;[8] it operates in a short-circuit regime. In fact, increasing the particle density by an order of magnitude requires a threefold increase in the current amplitude and an increase in the stored energy by an order of magnitude. In practical cases, this is impossible, since it requires an order of magnitude increase in the capacitance or three times the initial voltage.

Many authors[7,23] have shown in studies of the dependence of the cutoff current on the delay between the time the plasma is injected and the beginning of the feed to the inductive storage system that, besides the increase in the delay time, there is an increase in the cutoff current, and on trig-

gering, the plasma opening switch operates in a short-circuit regime.

In the proposed model an increase in the delay time leads to a simultaneous increase in R, the characteristic scale length of the inhomogeneity, and in the plasma density. The increase in R is caused by expansion of the injected plasma and the increase in density, by continuous injection of plasma opening switch plasma and its reflection from the electrodes. Equation (2) implies that this requires a rise in the cutoff current. When the delay time is increased further, R and n reach the values corresponding to a cutoff current greater than I_0, operation ceases, and a short-circuit regime is realized.

Operation can begin again after the passage of along (~ 100 μs) delay time, as observed experimentally.[23] The continuously injected gun plasma fills a large volume and becomes stable with respect to currents greater than I_0. On the other hand, it continually recombines and decays. The most rapid decay occurs at the edge of the plasma. In the center it is replaced for a long time by injection. After the plasma gun stops operating, recombination and decay reduce the volume occupied by the plasma. The characteristic scale length and density again begin to satisfy Eq. (2) and the plasma opening switch starts to operate.

The recombination and decay of the plasma require a certain time. This time may be determined by the experimentally observed[2] time shift between the operation of the plasma opening switch and the passing of the plasma gun current through zero in the course of its periodic variation. Typically, after the plasma gun current passes through zero the cutoff current initially falls off and then again increases and operation ceases until the gun current passes again through zero. The plasma opening switch operates right after the plasma gun current passes through zero, when R and n are essentially minimal. This picture is confirmed in an especially convincing manner by experiments with an aperiodic gun current. Thus, the proposed dynamic model for a plasma opening switch provides a qualitative explanation for the surprising dependence of the cutoff current on the plasma gun current waveform.[2]

The expansion of the plasma and its compression in pinches can begin with an increasing current near I_0. Compression leads to a drop in the characteristic scale length R of the inhomogeneity which ensures further compression, even for currents $I < I_0$, i.e., in the falling portion of the current. This experimentally observed fact[2] finds a natural explanation in the proposed model and cannot be explained at all in terms of the earlier model.[1]

Naturally, the question of the stability of compression in parallel pinches arises in the proposed model. Assuming that the initial inductance of the pinches is small, it is easy to show in the electrostatic approximation that two parallel pinches operate stably.

ESTIMATE OF THE SWITCHING TIME AND VOLTAGE

Because of the inhomogeneity of the plasma, the cylindrical geometry, reflection of the plasma from the electrodes, and the acceleration of the plasma along the storage system

by the nonuniform magnetic field, Eq. (2) cannot be satisfied simultaneously over the entire length of an inhomogeneity; this leads to the development of a noncylindrical Z-pinch which ends in sausages.[24] As noted above, the localization of high electric fields observed experimentally in plasma opening switch plasmas[20] can be a result of the development of sausages and serve as confirmation of their appearance.

Collapse of the plasma in pinches increases the inductance of the conducting plasma channels. High voltage pulses develop across the increasing inductances and these switch the current from the plasma opening switch storage system to the load. It has been shown[25] that for instability rise times much shorter than the current rise time, the voltages that are generated can be much higher than the initial capacitor voltage and are determined mainly by the rate of increase of the inductance of the pinch. In fact, under these assumptions we can regard the magnetic field as constant over the time the plasma opening switch operates and for a small pinch inductance L we obtain

$$L_0 \cdot I_0 = (L_0 + L)I,$$

which yields $U = -L_0 \dot{I} = \dot{L} I_0$. According to magnetohydrodynamics,[25] $\dot{L}$ increases without bound in a sausage and, along with $\dot{L}$, the voltage that is generated.

The snowplow model, which provides a satisfactory description of the initial phase of a Z-pinch and gives the right order of magnitude for the compression time,[26,27] yields $\tau = \pi c \rho^{1/2} R / I_0$ for the characteristic compression time and, therefore, the switching time. When n is changed from 10^{12} to 10^{14} cm^{-3} and the currents from 10^5 to 10^6 A, this expression yields a value of $10^{-8} - 10^{-7}$ s for τ. These switching times are in satisfactory agreement with experiment.[2,7]

The characteristic magnitude of the voltage when the instability develops is given by[25] $U_0 = V_0 \, T / 4\pi\tau$, where T is the period of the discharge in the ''hard'' circuit. For $\tau \approx 10^{-8}$ s this expression gives the characteristic magnitude of the voltages that develop during switching. For a plasma opening switch[7] with $T = 2.5 \times 10^{-7}$ s, the above formula gives $U_0 \approx 3 V_0$, and for microsecond plasma opening switches with $T = 1.5 \times 10^{-6}$ s, we obtain $U_0 \approx 10 V_0$. Usually a plasma opening switch operates on a diode load and the maximum voltages are limited by breakdown of the diode, so they do not reach these levels. When a plasma opening switch operates on an inductive load, the generated voltages are determined by the inductance and equal[25]

$$L_n \dot{I} \approx \frac{L_n}{L_0 + L_n} \cdot \frac{L_0 \cdot I_0}{\tau} \approx 2 V_0,$$

which is also in satisfactory agreement with experiment.[2]

In designs with a relative gap less than 1 and a large number of plasma guns, the injected plasma is more uniform spatially and may be stabilized with respect to pinches and sausages. In this case the picture is very different from that described above. Let us estimate the velocity of the plasma from experimental data for the KALIF project.[28] We have chosen this experiment because its parameters have been disclosed in greatest detail.

In terms of an energy approach, assuming that the energy in the storage system initially goes into the energy of the magnetic field in the inductance and into the kinetic energy of the motion of the entire plasma, we obtain $V \approx 4 \times 10^8$ cm/s, while the acceleration dynamics for the entire plasma as a whole implies $V \approx 3 \times 10^8$ cm/s. The snowplow model gives a maximum velocity of $V \approx 3.5 \times 10^8$ cm/s. For such a velocity the electric field strength at the central electrode is $\approx 3 \times 10^6$ V/cm, while the voltage $L \dot{I}_0 \approx 2 \times 10^6$ V. In the best shots[28] the maximum voltage measured by nuclear techniques is 3.8×10^6 V. The agreement between these voltages and the experimental measurements is not by chance. The acceleration of the plasma in these designs is a consequence of the development of force instabilities that cause a rise in the inductance.

CONCLUSION

The dynamical model for plasma opening switches based on the appearance and development of force instabilities proposed here provides a natural explanation of the existence of a critical current (the cutoff current) and of its dependence on the particle density, the number of plasma guns, the delay between the time plasma is injected and the time the storage system is filled, and the current waveform. It also explains the operation of plasma opening switches when the polarity of the electrodes is changed, their triggering on a falling current, the speed of penetration of the magnetic field along the storage system, and the generation of charged particles. The proposed model can be used to obtain an estimate of the operating time and the magnitude of the voltage that is generated.

The author hopes that this dynamic model for plasma opening switches holds true and notes that in the form presented here it is still too rough and requires detailed development.

[1] P. F. Ottinger, S. A. Goldstein, and R. A. Meger, J. Appl. Phys. **56**, 774 (1984).

[2] E. N. Abdullin, G. P. Bazhenov, A. A. Kim *et al.*, Fiz. Plazmy **12**, 1260 (1986) [Sov. J. Plasma Phys. **12**, 728 (1986)].

[3] G. V. Ivankov, Zh. Tekh. Fiz. **61**(5), 46 (1991) [Sov. Phys. Tech. Phys. **36**, 524 (1991)].

[4] A. V. Gordeev, A. V. Grechikha, A. V. Gulin *et al.*, Fiz. Plazmy 17, 650 (1991) [Sov. J. Plasma Phys. **17**, 381 (1991)].

[5] P. V. Sasorov, JETP Lett. **56**, 599 (1992)].

[6] G. I. Dolgachev, L. P. Zakatov, Yu. G. Kalinin *et al.*, Fiz. Plazmy **22**, 1017 (1996) [Plasma Phys. **22**, 921 (1996)].

[7] B. V. Weber, R. J. Commisso, G. Cooperstein *et al.*, IEEE Trans. Plasma Sci. **PS-15**, 635 (1987).

[8] Yu. P. Golovanov, G. I. Dolgachev, L. P. Zakatov *et al.*, Fiz. Plazmy **14**, 880 (1988) [Sov. J. Plasma Phys. **14**, 519 (1988)].

[9] B. M. Bystritskaya, A. N. Didenko, S. N. Volkov *et al.*, Fiz. Plazmy **12**, 1178 (1986) [Sov. J. Plasma Phys. **12**, 679 (1986)].

[10] É. N. Abdullin, G. P. Erokhin, V. M. Kiselev *et al.*, *Abstracts of the VII All-Union Symp. on High-Current Electronics* [in Russian], Tomsk (1988), Part III, pp. 49–51.

[11] L. A. Artsimovich and R. Z. Sagdeev, *Plasma Physics for Physicists* [in Russian], Atomizdat, Moscow (1979), 320 pp.

[12] A. L. Velikovich and M. A. Liberman, *Physics of Shock Waves in Gases and Plasmas* [in Russian], Nauka, Moscow (1987), 296 pp.

[13] D. D. Hinshilwood, J. R. Boller, R. J. Commisso *et al.*, IEEE Trans. Plasma Sci. **PS-15**, 564 (1987).

[14] É. N. Abdullin, G. P. Bazhenov, V. M. Bystritskiĭ *et al.*, Fiz. Plazmy **13**, 1027 (1987) [Sov. J. Plasma Phys. **13**, 589 (1987)].

[15] V. A. Kokshenev, *Abstracts of the VII All-Union Symposium on High-Current Electronics* [in Russian], Tomsk (1988), Part III, pp. 16–18.

[16] B. B. Kadomtsev, *Nonlinear Waves* [in Russian], Nauka, Moscow (1979), pp. 131–150.

[17] A. F. Aleksandrov and A. A. Rukhadze, *Physics of High-Current Electrical Discharge Light Sources* [in Russian], Atomizdat, Moscow (1976), 184 pp.

[18] G. A. Mesyats, A. N. Didenko, É. N. Abdullin *et al.*, Dokl. Akad. Nauk **289**, 84 (1986).

[19] L. A. Artsimovich, *Controlled Thermonuclear Reactions* [in Russian], GIFML, Moscow (1961), 468 pp.

[20] Yu. P. Golovanov, G. I. Dolgachev, L. P. Zakatov *et al.*, Fiz. Plazmy 17, 799 (1991) [Sov. J. Plasma Phys. **17**, 466 (1991)].

[21] B. M. Koval'chuk, V. A. Kokshenev, and F. I. Fursov, *Abstracts of the VI All-Union Symposium on High-Current Electronics* [in Russian], Tomsk (1986), Part II, pp. 139-141.

[22] A. G. Mozgovoĭ, *Abstracts of the VI All-Union Symposium on High-Current Electronics* [in Russian], Tomsk(1988), Part III, pp. 7–9.

[23] G. A. Mesyats, É. N. Abdullin, A. N. Bystrikov *et al.*, Fiz. Plazmy **11**, 109 (1985) [Sov. J. Plasma Phys. **11**, 66 (1985)].

[24] V. F. D'yachenko and V. S. Imshennik, in *Reviews of Plasma Physics* [in Russian], Vol. 8, Atomizdat, Moscow (1974), pp. 164–246.

[25] P. I. Zubkov, Zh. Prikl. Mekh. Tekh. Fiz., **4**, 24–41 (1993).

[26] P. I. Zubkov, S. M. Ischenko, and K. A. Ten, *Abstracts of the VII International Conference on the Generation of Megagauss Magnetic Fields and Related Experiments* [in Russian], Sarov (1996), pp. 41–42.

[27] P. I. Zubkov, L. A. Luk'yanchikov, and K. A. Ten, Zh. Prikl. Mekh. Tekh. Fiz., **2**, 182–187 (1990).

[28] H. Bluhm, K. Bohnel, P. Hoppe *et al.*, IEEE Trans. Appl. Supercond. **PS-15**, 654 (1987).

Translated by D. H. McNeill

Approximate similarity of electrophysical and kinematic processes during pulsed corona discharges in strong electrolytes

V. V. Shamko, E. V. Krivitskiĭ, and V. V. Kucherenko

*Institute of Pulsed Processes and Technologies, National Academy of Sciences of Ukraine,
327018 Nikolaev, Ukraine*
(Submitted December 26, 1997; resubmitted October 6, 1998)
Zh. Tekh. Fiz. **69**, 30–34 (May 1999)

Based on a theoretical method of dimensional and similarity analysis and on published experimental data on the dynamics of corona discharges in strong electrolytes, dimensionless numbers are found which make manifest the approximate similarity of electrophysical and hydro/gasdynamic phenomena. A physical experiment is carried out, and generalized time dependences are obtained for the current, voltage across the interelectrode gap, and plasma radius in terms of these similarity numbers. © *1999 American Institute of Physics.*
[S1063-7842(99)00605-4]

A pulsed corona in water[1,2] is a developed system of leaders with a clustered spatial orientation. This orientation develops in discharge gaps with nonuniform and highly non-uniform fields and is characterized by the fact that none of the leaders growing into the depth of the gap reaches the opposite electrode. This is the significant difference between a pulsed corona and a "linear" underwater spark discharge.[3]

In the case of a discharge that is not completed by break-down of the interelectrode gap, the entire discharge current flows through the boundary between the leader system plasma and the liquid medium. Because of the comparatively small plasma–medium contact surface area and the high resistance of the liquid layer between the clustered system of leaders and the counterelectrode, the discharge current is extremely low, so there is little heating of the plasma in the leaders. Since the conductivity of the plasma is clearly higher than that of the medium, a substantial fraction of the energy is released in the surrounding liquid, rather than in the corona itself.

When the external conditions are changed, especially the specific electrical conductivity σ_0 of the medium, the number of leaders increases, their diameter at the base becomes larger, and, for certain electrically conducting media and field strengths,[1,2] they cover the entire exposed surface of the electrode and are converted into a continuous plasma formation (plasmoid). When threshold values of the conductivity of the medium are reached, the leaders merge with one another from the very beginning of the discharge and the plasmoid fully reproduces the shape of the electrode tip. As a result, the energy of Joule heating is released in four regions: the plasmoid, the plasma–liquid transition layer, the volume of the conducting liquid, and equivalent active elements of the discharge circuit. The experimental data accumulated up to now[1,2,4] on the phenomenology of this effect show that the characteristics of corona discharges depend significantly on the external controllable parameters of the discharge circuit, the conducting medium, and the electrode geometry.

The close attention given to corona discharges in recent years is related to the experimentally established fact[2] that,

under certain conditions, they have hydrodynamic instabili-ties of rather high intensity comparable, at least, to the per-turbations generated by underwater spark discharges. This kind of hydrodynamic behavior is intrinsic to corona dis-charges in strong electrolytes, which typically have a con-tinuous plasmoid whose shape can be controlled, so that a pressure distribution with a given configuration can be cre-ated in the liquid. In this paper we shall discuss just this sort of discharge.

Although some first attempts have been made to develop an electrodynamic model of corona discharges in strong electrolytes[2,5] and interpret them hydrodynamically,[6] and al-though individual aspects of the phenomenon have been modeled somewhat satisfactorily, there is still not enough factual material in this area to create a correct mathematical and physical model of the phenomenon. Thus, for system-atizing the available experimental data and comparing them, given the differences in the experimental conditions, it seemed appropriate to generalize them on the basis of a the-oretical method of dimensional and similarity analysis.[7]

The basis for seeking the structure of dimensionless numbers is a set of dimensional physical quantities which determine the desired dynamical characteristics of the dis-charge with sufficient accuracy. In order to establish a set of dimensional parameters, it is necessary to assess the initial knowledge about the phenomenon under study. This task was performed by making a careful analysis of the experi-mental papers[1,2,4] containing the most systematic experimen-tal data on corona discharges having a continuous plasmoid.

In studies of electrophysical phenomena, the dependent variables are the current $i=i(t,\{\beta_i\})$ and voltage $U=U(t,\{\beta_U\})$ within the volume and the hydrodynamic variables are the radius of the plasmoid and bubble $a=a(t,\{\beta_a\})$, the pressures $P_a=P_a(t,\{\beta_P\})$ in them, and the pressures in the primary and secondary compression waves $P=P(t,r,\{\beta_r\})$. The independent variables t, r, and $\{\beta_j\}$ in the arguments are the time, spatial coordinate, and a set $\{\dots\}$ of dimensional and dimensionless parameters important for the development of one or another aspect of the phenomenon.

On analyzing the systematic experimental studies available today on corona discharges in strong electrolytes,[1,2,4] one can state that their electrophysical characteristics are most strongly affected by the electrical conductivity of the liquid electrolyte. For the initial field strengths $E_0 = U_0/r_0 = 10^6 - 10^8$ V/m that have been studied, the threshold conductivity σ_* at which the discharge changes from a branching corona into a continuous plasmoid lies within the interval $\sigma_0 = 1.6 - 3.3$ S/m. The value of σ_* is determined by the field strength E_0 and decreases as the latter is raised.[1] On passing through σ_* the voltage curves $U(t)$ change from those typical of an underwater spark[8] to bell-shaped curves, close in shape to that of the current, and their maxima $U_{m1} \approx (0.6 - 0.9)U_0$ lie below the charge U_0 on the capacitor bank. On passing through $\sigma_0 \approx 6$ S/m, the discharge changes from aperiodic to oscillatory. As σ_0 increases, the amplitudes of the current and power increase, while the duration of the first half wave of the current (voltage) decreases. Thus, for $3.3 \leqslant \sigma_0 \leqslant 20$ S/m, the power amplitude N_{m1} over the interval $0.4 \leqslant N_{m1}/(U_0^2\sqrt{L/C}) \leqslant 1$ is very much greater than for an underwater spark.[8] The qualitative variation in the electrophysical characteristics with E_0 is analogous to that described above for the conductivity, and their reaction to changes in the electrolyte temperature T_0 is completely equivalent to their reaction to a corresponding change in $\sigma_0(T_0)$. The electrophysical characteristics are insensitive to the hydrostatic pressure P_0 of the medium, at least within the range from 0.1 to 30 MPa.[4]

The hydro/gasdynamic characteristics of corona discharges have some distinctive features compared to the electrophysical characteristics. In particular, they do not react accordingly to the corresponding changes in the electrical conductivity and temperature of the medium. For example, the amplitude of the secondary compression wave increases as σ_0 increases but decreases with increasing T_0. On the other hand, the amplitude of the primary wave, although it also increases with rising temperature, has an extremum in the region of $\sigma_0 \approx 8 - 12$ S/m,[4] depending on σ_0. After $\sigma_0 > 1$ S/m, the maximum radius a_{max} of the vapor–gas cavity and the period t_1 of its first pulsation essentially do not react to changes in σ_0 (although a noticeable dependence $t_1(\sigma_0)$ has been observed at elevated hydrostatic pressures, at least for $P_0 = 30$ MPa),[4] while the depend very strongly on T_0, even for $\sigma_0(T_0) > 1$ S/m. The amplitude of the first wave is affected significantly by the radius r_D of the tip of the positive electrode and the length l of the interelectrode gap (when l is short),[4] which increase P_{m1} as they become smaller. The hydro/gasdynamic parameters are found to have a significant dependence on the hydrostatic pressure of the medium, becoming quite substantial for the cavity parameters. There are, as yet, no experimental data on the pressure in the plasmoid and, as far as its kinetics are concerned, we can only say that as σ_0 increases, the radius of the plasmoid decrease as the energy release comes to an end.[2]

It should also be noted that there are, as yet, no systematic data on the effect of such circuit parameters as the capacitance C, inductance L, and equivalent short-circuit resistance R_k of the discharge circuit on the discharge dynamics. The capacitance and inductance affect the discharge regime depending on the relationship between the characteristic impedance of the circuit and the load resistance, which includes the total resistance of the plasmoid, R_a, and electrolyte layer, R_e, averaged over the discharge time, and R_k. For $R_a + R_e + R_k \geqslant 2\sqrt{L/C}$ the discharge will be aperiodic; otherwise, it will be oscillatory. Here the role of the initial conductivity of the electrolyte, which was mentioned above in connection with the oscillatory character of the discharge, reduces to a change in the resistance R_l. Because of the above mentioned features, and also because these circuit elements (C, L, R_k) must be formed during the discharge, they will be included in the system of parameters controlling this phenomenon.

Since one of the elements of the object under study is a plasmoid, by analogy with underwater sparks[8] we shall use the spark constant A as a characteristic of the discharge plasma. A is the coefficient of proportionality between the pressure and electrical conductivity of the plasma, and the effective adiabatic index γ of the plasma. The initial density ρ_0 of an electrolyte based on $H_2O + NaCl$ varies as the concentration of the solution is increased up to saturation at 20%. From a hydrodynamic standpoint, this is a rather significant factor, so the parameter ρ_0 must be included among the system of control parameters in a study of the hydro/gasdynamic phenomena.

Here we shall only consider those discharge regimes in which a continuous plasmoid is formed. In accordance with Refs. 1 and 2, this places a lower bound on the conductivity of the electrolyte. Then, at least for $E_0 \leqslant 10^8$ V/m,[2] the time to ignite a corona and the amount of electrical energy expended in this stage are incomparably smaller than the characteristic discharge time $\sqrt{L/C}$ and the energy stored in the capacitor bank, respectively. The ignition time is observed to decrease significantly with rising E_0. Therefore, the unknown dependent parameters of the corona discharge will not be sensitive to the factors controlling the discharge ignition phase. As a result, we can include the capacitor bank charging voltage and the radius of the positive electrode rod in the system of control parameters in place of the initially unknown ignition voltage and the initial radius of the plasmoid, respectively.

During a corona discharge the electrolyte is heated at a constant external pressure, and a temperature distribution develops within it owing to resistive dissipation within the volume. When the effects of molecular heat conduction are small,[2] the density of the internal heat sources is proportional to the rate of change of the temperature with a coefficient of proportionality c_p, which represents the specific heat at constant pressure. Therefore, the parameter c_p should also be included in the system of control parameters.

To summarize, we can distinguish the independent physical quantities σ_0, U_0, C, L, r_0, l, ρ_0, T_0, A, P_0, R_k, c_p, t, r, and γ which determine the evolution of a corona discharge in strong electrolytes and, in accordance with the bulk of the original experimental data and physical considerations, adequately characterize the electrophysical and hydro/gasdynamic phenomena within it. Therefore, with 14 independent dimensional variables and the 5 primary dimensions (kg, m, s, A, and K), according to the π theoremof

dimensional analysis we shall have 9 independent dimensionless numbers. If we choose the quantities U_0, C, L, T_0, and r_0 as the variables with independent dimensions, then the independent dimensionless numbers can be written explicitly as follows:

$$\Pi_A = A r_0^2/(U_0^2 \sqrt{LC}), \quad \Pi_\sigma = \sigma_0 r_0 \sqrt{L/C},$$

$$\Pi_R = R_k \sqrt{C/L}, \quad \Pi_c = c_p T_0 LC/r_0^2,$$

$$\Pi_\rho = \rho_0 r_0^5/(LC^2 U_0^2), \quad \Pi_{P_0} = P_0 r_0^3/(CU_0^2),$$

$$\Pi_l = l/r_0, \quad \Pi_r = r/r_0, \quad \Pi_t = t/\sqrt{LC}, \quad \Pi_\gamma = \gamma. \quad (1)$$

For the dependent variables i, U, P, and a of interest to us, four additional dependent dimensionless numbers can be written:

$$\Pi_i = i\sqrt{L/C}/U_0, \quad \Pi_U = U/U_0,$$

$$\Pi_P = P r_0^3/(CU_0^2), \quad \Pi_a = a/r_0. \quad (2)$$

The number Π_A, which corresponds qualitatively to the well known channel number Π_k for an underwater spark[9] and characterizes the complex action of the electrical parameters of the discharge circuit and plasmoid on the medium, is not entirely convenient, for example, in analyzing the current–voltage–time characteristics over the entire discharge interval. The following combination of the numbers Π_A and Π_σ is more convenient for these purposes:

$$\Pi_{A\sigma} = \Pi_A \cdot \Pi_\sigma = A\sigma_0 r_0^3/(CU_0^2), \quad (3)$$

which is the ratio of the initial integrated energy density in the volume to the electrical energy density of the source. Because of the low energy densities in the plasma and electrolyte compared to CU_0^2/r_0^3, it is appropriate to replace the number Π_P (and Π_{P_0}) by a combination of the numbers Π_P, Π_σ, and Π_A:

$$\Pi_P' = \Pi_P \cdot (\Pi_\sigma \cdot \Pi_A)^{-1} = P/(A\sigma_0), \quad \Pi_{P_0}' = P_0/(A\sigma_0). \quad (4)$$

It is more appropriate to rewrite the dimensionless number Π_σ in the form

$$\Pi_\sigma' = \Pi_\sigma^{-1} = (\sigma_0 r_0)^{-1}/\sqrt{L/C}, \quad (5)$$

and then the dimensionless numbers Π_σ' and Π_R will represent the ratios of the resistances of the layer of electrolyte (for $r_0 \ll l$) and the circuit to the impedance, respectively.

Therefore, given Eqs. (3)-(5), the above statement of the problem will correspond to generalized solutions in the form of the dimensionless functions

$$\Pi_i = F_i(\Pi_t, \Pi_{A\sigma}, \Pi_\sigma', \Pi_R, \Pi_c, \Pi_\gamma),$$

$$\Pi_U = F_U(\Pi_t, \Pi_{A\sigma}, \Pi_\sigma', \Pi_R, \Pi_c, \Pi_\gamma),$$

$$\Pi_{a1} = F_{a1}(\Pi_t, \Pi_{A\sigma}, \Pi_\sigma', \Pi_R, \Pi_c, \Pi_\gamma),$$

$$\Pi_P' = F_P(\Pi_t, \Pi_R, \Pi_{A\sigma}, \Pi_c, \Pi_\rho, \Pi_\sigma', \Pi_l, \Pi_\gamma),$$

$$\Pi_{a2} = F_{a2}(\Pi_t, \Pi_{A\sigma}, \Pi_c, \Pi_\rho, \Pi_{P_0}, \Pi_\gamma). \quad (6)$$

Here Π_{a1} corresponds to Π_a for the plasmoid and Π_{a2} corresponds to that for the cavity. Obtaining dimensionless dependences of the form (6) exhausts the possibilities of dimensional and similarity analysis. Approximate expressions for the functions F_j in explicit form must be sought by generalizing the experimental data. Usually this is done by representing the functions F_j in the form of power law functions of their arguments and fitting the exponents to experimental data. Although this approach does ensure fairly good accuracy within the range of variation of the parameters in the experiments, the resulting formulas essentially do not work beyond these limits. Here it should also be noted that the possibility of modeling phenomena when more than two independent dimensionless numbers are involved (except for the dimensionless coordinates and time) is extremely problematical. Thus, it is more appropriate to begin with physical considerations and isolate the main dimensionless number (or several of them) and then to show, by generalizing the experimental data, that major changes in the secondary dimensionless numbers do not change the picture of the process.

Thus, for discharges in media with a single chemical constituent, Π_γ drops out of the set of dimensionless numbers, as it is derived only from the physical properties of the medium and is constant. Π_l has an effect on the discharge dynamics only when $\Pi_l < 1$, so that for corona discharges with a continuous plasmoid, for which $\Pi_l \gg 1$ (Ref. 2), it can be neglected. The dimensionless numbers Π_R and Π_ρ are much smaller than unity, so we shall neglect them as well, since $\Pi_{A\sigma}$ and Π_σ' are of the order of unity. As result, instead of the dimensionless functions (6), we can write down their approximate analogs

$$\Pi_i = \Psi_i(\Pi_t, \Pi_{A\sigma}, \Pi_\sigma'),$$

$$\Pi_U = \Psi_U(\Pi_t, \Pi_{A\sigma}, \Pi_\sigma'),$$

$$\Pi_{a1} = \Psi_{a1}(\Pi_t, \Pi_{A\sigma}, \Pi_\sigma'),$$

$$\Pi_P' = \Psi_P(\Pi_t, \Pi_r, \Pi_{A\sigma}, \Pi_\sigma', \Pi_c),$$

$$\Pi_{a2} = \Psi_{a2}(\Pi_t, \Pi_{A\sigma}, \Pi_c, \Pi_{P_0}). \quad (7)$$

It is easy to see that the structure of the generalized functions for the electrophysical characteristics and radius of the plasmoid have been greatly simplified and they can now be subjected to a generalized experimental test. To do this we have done an experiment with ranges for the dimensional physical quantities $U_0 = 15$–47 kV, $C = 0.5$–12 μF, $L = 2.4$–3.8 μH, $\sigma_0 = 3$–22 S/m, and $r_0 = 0.5$–5 mm such that the numbers $\Pi_{A\sigma}$ and Π_σ' remained constant. Ten discharges were produced for each set of fixed values of these quantities, and the discharge current $i(t)$, voltage $U(t)$ across the discharge gap, plasmoid radius $a(t)$, and amplitude P_{m1} of the primary compression wave were recorded in each. The data were then normalized to the corresponding scales $i_M = U_0\sqrt{C/L}$, $U_M = U_0$, $a_M = r_0$, $P_M = A\sigma_0$, and $t_M = \sqrt{LC}$. The measurement technique described in Ref. 2 was used, along with the experimental data obtained there.

Figure 1 shows dimensionless plots of the current and voltage as functions of time for fixed similarity numbers ($\Pi_{A\sigma} = 2.5 \cdot 10^{-6}$ and $\Pi_\sigma' = 29.4$) (the vertical bars on the curves denote the 90% confidence intervals). It is easy to see that the number Π_U (Fig. 1b) is uniquely determined by the
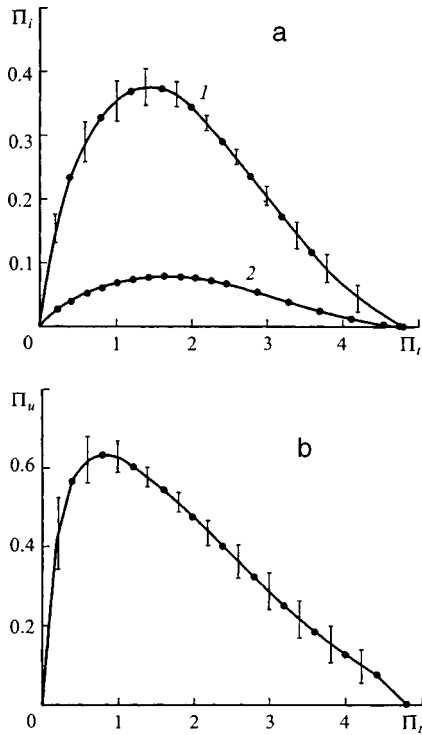
FIG. 1. Generalized time variations in the current (a) and voltage (b) for $\beta = 0.35$.

two numbers $\Pi_{A\sigma}$ and Π_{σ}', while Π_i (Fig. 1a) splits into two regions: in the first (curve *1*) there are seven current regimes with a maximum of $\Pi_{i\max} = 0.38$ and in the second (curve *2*), three regimes with $\Pi_{i\max} = 0.08$. Therefore, in terms of the approximate similarity considered here, the current charac-
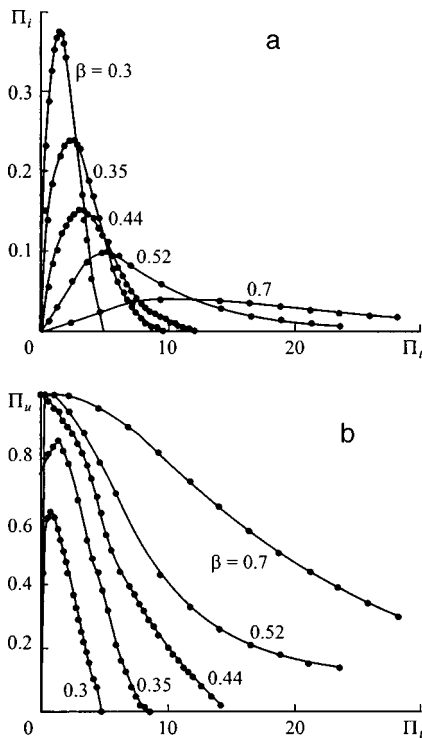

FIG. 2. Normalized plots of the current (a) and voltage (b) as functions of time and of the dimensionless number β.
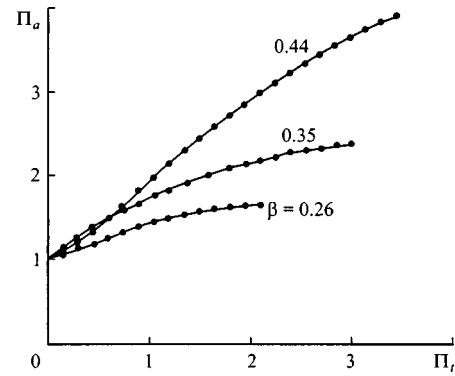
FIG. 3. Normalized plots of the radius of a continuous plasmoid as a function of time and the dimensionless number β.

teristics are subject to a scaling effect. Since the distinguishing feature of these two classes of electrophysical effects is a difference in the characteristic times (for the first series of curves $t_M \leq 4.2$ μs, and for the second, $t_M > 4.2$ μs), it is possible that the current scaling effect (as observed before[8] for underwater spark discharges) is related to some threshold time beyond which the approximate similarity breaks down. Further generalization of the experimental data was carried out only for characteristic times $t_M \leq 4.2$ μs.

The extremely wide range of the dimensionless number $\Pi_{A\sigma} = 10^{-10} - 1$ makes it an inconvenient parameter, at least for classifying the discharge regimes. Thus, instead of $\Pi_{A\sigma}$ we have constructed a combination of the numbers $\Pi_{A\sigma}$ and Π_{σ}' that is more convenient in terms of its numerical range,

$$\beta = 0.1[\Pi_{\sigma}'/(4\pi^2 \Pi_{A\sigma})]^{0.1}, \qquad (8)$$

which is qualitatively similar to the dimensionless number[9] η for underwater spark discharges. For low β the discharge is oscillatory, and as $\beta \to 1$ it becomes aperiodic (Fig. 2). The number β is more important than Π_{σ}' in its effect on the generalized time variations in the current, voltage, and plasmoid (Figs. 2 and 3).

The normalized distribution of the maximum pressure in the primary compression wave $\pi_{P_{m1}}$ (Π_r) for fixed $\beta = 0.35$ is illustrated by Table I (n is the number of the discharge regime and $\beta = $ idem).

In the first approximation this distribution has an acoustic dependence.

Given that the resistance of the electrolyte layer will prevail over R_a and R_k, we conclude that Π_{σ}' will be the second principal dimensionless number of approximate similarity of the electrophysical and kinematic characteristics of a corona discharge in strong electrolytes. However, as we have done for $\Pi_{A\sigma}$, it is better to use the construct

TABLE I.

Π_r	$\Pi_{P_{m1}}$	n
33	15.2	2
50	11.8	3
100	5.1	1

$$\delta = (\Pi_\rho \cdot \Pi'_\sigma)^{-1} = \sigma_0 E_0^2 / [\rho_0 r_0^2 / (LC)^{3/2}], \qquad (9)$$

which is a measure of the ratio of the energy flux of the electric field to the energy flux of the dynamic head of the liquid medium. We believe that the similarity number δ, on which the corona discharge characteristics studied here are less dependent, does determine the threshold for the transition of a branched pulsed corona into a discharge with a continuous plasmoid. A preliminary estimate for this threshold is $\delta_* = 0.2$. When $\delta > \delta_*$ the discharge is a continuous plasmoid.

[1] V. Ya. Ushakov, *Pulsed Electrical Breakdown of Liquids* [in Russian], TGU, Tomsk (1975), 256 pp.

[2] L. Z. Boguslavskiĭ, V. V. Kucherenko, and E. V. Krivitskiĭ, IIPT Preprint No. 22 [in Russian], Institute of Pulsed Processes and Technologies, Nikolaev (1993), 41 pp.

[3] E. V. Krivitskiĭ, Zh. Tekh. Fiz. **61**(1), 9 (1991)[Sov. Phys. Tech. Phys. **36**, 4 (1991)].

[4] V. Zh. Glushchenko, V. G. Zhekul, and L. P. Trofimova, in *Theory, Experiment, and Practice of Electrical Discharge Technologies* [in Russian], Naukova Dumka, Kiev (1995), No. 2, pp. 4–7.

[5] V. A. Pozdeev, N. M. Beskaravaĭnyĭ, and V. K. Sholom, Élektron. Obrab. Mater., **3**, 33–36 (1990).

[6] Zh. N. Ishchenko, in *Theory, Experiment, and Practice of Electrical Discharge Technologies* [in Russian], Naukova Dumka, Kiev (1995), No. 2, pp. 55–59.

[7] S. S. Kutateladze, *Similarity Analysis and PhysicalModels* [in Russian], Nauka, Novosibirsk (1986), 296 pp.

[8] E. V. Krivitskiĭ and V. V. Shamko, *TransientProcesses in High Voltage Discharges in Water* [in Russian], Naukova Dumka, Kiev (1979), 208 pp.

[9] E. V. Krivitskiĭ and V. V. Shamko, Zh. Tekh. Fiz. **42**, 83 (1972) [Sov. Phys. Tech. Phys. **17**, 62 (1972)].

Translated by D. H. McNeill

Periodic injector with a porous pellet mold for introducing fuel into plasmas

I. V. Vinyar

St. Petersburg State Technical University, 195251 St. Petersburg, Russia

A ten-barrel pneumatic injector for periodic introduction of fuel pellets into steady-state thermonuclear experiments has been developed. Solid hydrogen pellets with diameters of 2.7 mm and lengths of 3 to 4 mm are formed in each barrel at a rate of 0.1–0.2 Hz with periodic pulsed heating of a porous insert, which is continuously filled with hydrogen and cooled with liquid helium, and accelerated to 1–1.2 km/s. © *1999 American Institute of Physics.*
[S1063-7842(99)00705-9]

INTRODUCTION

Thermonuclear reactors must be equipped with systems for delivering fuel in order to operate in a steady state. Besides gas puffing, it has been proposed that pellets of solid hydrogen isotopes be introduced into a reactor, as they penetrate deeper into the plasma and have a positive effect on its parameters.[1] Here the key problem is to develop a reliable injector capable of delivering fuel pellets deep into the plasma in an uninterrupted periodic sequence at a rate of 1–10 Hz. Three methods have already been proposed for the continuous formation and periodic injection of pellets. In the first, pellets are created by extrusion: either by alternating operation of 2–3 piston extruders,[2,3] using an extruder with a ''gas piston,''[4] or using a screw extruder.[5] in the second method, fuel gas is continuously frozen at the rim of a rotating disk and the resulting ice is cut to form pellets.[6] In the recently proposed, third method, pellets are produced by melting a small volume of initially frozen fuel which then penetrates into a barrel through a porous material and is frozen there.[7] Pellet injection at a rate of about 0.05 Hz has been demonstrated with a single barrel injector using this technology and increased to 0.1 Hz by shooting in series of 3–5 pellets each.[8] The multibarrel injector described in this paper was created in order to raise the injection rate above 0.1 Hz.

DESIGN AND OPERATING PRINCIPLE OF THE INJECTOR

The injector consists of ten single-stage light-gas guns with a pellet mold, a diagnostic chamber equipped with a helium–neon laser and photodetector for recording the transit time of the pellets, a vacuum system, an electronic power supply rack including a logic controller. When the injector is attached to a thermonuclear machine, a differential pumping system with guide tube for passage of the pellets and a microwave device for measuring their mass is mounted between the two. Figure 1 is a sketch of the injector design. The light-gas guns are located in two rows of five. Each gun is equipped with a pulsed valve *2* for introducing the driver gas into the barrel, a valve *5* for introducing the fuel gas into its independent porous mold *1* and a gate valve *9*. The valves for the driver and fuel gases are mounted in the upper part of

the front flange of the cryogenic vessel *8*, inside of which lie the porous pellet molds and barrels *6*, which are protected by a heat shield *7*. Ten low-heat-conducting valves *3* are attached to the bottom of the front flange; these control the flow of vapor from the liquid helium through the heat exchanges on the ten porous molds. Four vacuum-tight feed throughs for the leads from the vessel for the heaters and temperature sensors are also located on the front flange. The gate valves are made up of two identical aluminum housings with five in a row. The gate consists of a steel plate with an aperture, 4 mm in diameter, for the pellets to pass through, compressed between two Teflon gaskets. The plate is moved by electromagnets and in 3–5 ms it hermetically closes off (opens) the barrel volume from the diagnostic chamber *10*. The latter is hermetically joined to the two housings of the gate valve and has two pairs of windows for monitoring the flight of the pellets by a photodetector illuminated by a laser beam and for video recording them in flight under the light of a nanosecond spark. The diagnostic chamber can be rotated and attached in any position. Its flange, which is located opposite to the barrels, is made of polymethyl methacrylate. A beam of light could be directed through it into the barrel so it was easy, with the valve for the driver gas off to the side, to observe the formation of a pellet directly in the barrel.

The fuel gas is fed through the valves *5* into each porous mold from a collector *4* equipped with a manometer. The driver gas enters from a tank with a regulator directly into the ten pulsed electromagnetic pilot valves with an internal volume of 6 cm^3 and an opening/closing time of about 1 ms. The cryogenic and diagnostic chambers are equipped with vacuum probes and are pumped by turbomolecular and roughing pumps, respectively.

The major element of the injector design which determines the injection rate is the pellet mold. Figure 2 is a sketch of one of the porous molds. It is made from a block of high conductivity copper *1* into which a series of copper grids with cell sizes smaller than 0.07 mm have been impressed. Apertures equal to the inner diameter of a barrel have been cut through the grids along one axis. A barrel *3* and pipe *6* are soldered to the opposite sides of the resulting porous insert *2* for delivery of the driver gas. A heat ex-
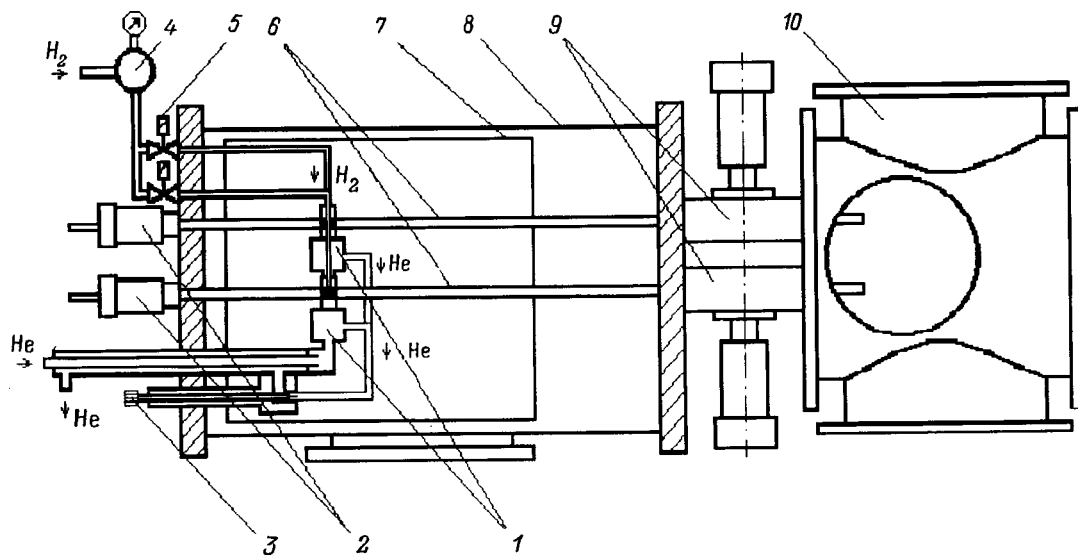
FIG. 1. A sketch of a multibarrel injector with a porous mold for fuel pellets.

changer channel *5* has been cut in the lower part of the mold and in the upper part a hole has been cut to join the porous insert to a pipe for the fuel gas feed. A temperature probe *4* is fastened to the middle portion of the mold and a manganin wire with a resistance of 20 Ω was wound around the upper portion as a heater *8*.

During operation of the injector, vapor from liquid helium constantly circulates through the heat exchanger and cools the porous insert. During a shot, the driver gas flows along a barrel, pushes a pellet *7* out of it, and heats the porous insert with the fuel frozen in it. The fuel (hydrogen isotopes with triple points ranging from 14 to 20.6 K) melts in the pores and penetrates into the barrel. There it freezes to form a new pellet. The driver gas pushes it out of the barrel during a regular shot and the cycle is repeated.
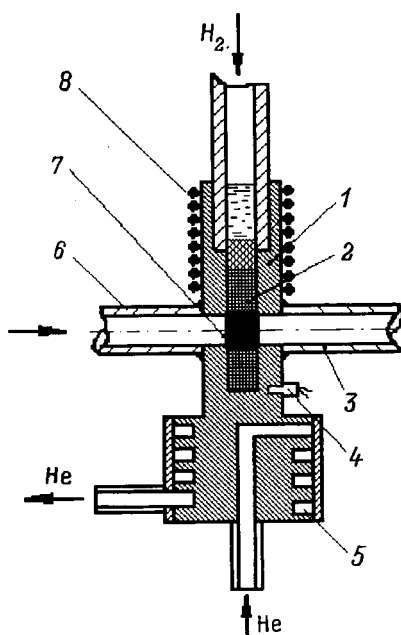
FIG. 2. Sketch of a porous pellet mold.

EXPERIMENTAL RESULTS

The purpose of the first stage of the experiments was to determine the minimum time for pellet formation. In this connection, tests were made with different mold designs, but with firing from a single barrel. Here we present the results of some experiments with hydrogen pellets, since no significant differences in the duration of these processes were observed in experiments with deuterium.

Two modes of pellet formation were realized during these tests. In the first, solid hydrogen was melted in the pores of the mold and the liquid droplet was frozen in the barrel. In the second, solid hydrogen was heated in the pores to temperatures below the melting (about 13 K) and the plastic hydrogen was extruded through the pores into the barrel and then cooled. The sequences of working pulses generated by the logic controller during the pellet injection cycles for these two modes are shown in Fig. 3. In both modes the injection cycle began with a pulse which opened the driver gas valve. 10 ms after the valve was opened for 0.4–0.8 ms, the 10-watt heater was turned on and delivered 4–8 J of heat to the mold. In the extrusion mode, hydrogen at a temperature of 13 K in the pipe above the mold moved, under a gas pressure of 3 MPa, into the barrel in the form of thin plastic filaments over a time of less than 0.5 s. This process was recorded by the video camera. Right after that, for 3–4 s the solid hydrogen was cooled in the barrel and the pellet formed was again fired by a pulse of gas from the valve (Fig. 3a). In this mode, the gate valve and hydrogen feed valve were always open. The process of forming and injecting the pellets is extremely simple and was performed by the driver gas feed valve and the heater. In the hydrogen melting mode (Fig. 3b), after the gas flowed out the barrel, the gate valve was closed for 80 ms. Solid hydrogen was melted in the pores and the liquid flowed into the barrel under the gas pressure of 0.1 MPa in the pipe above the mold. Part of the hydrogen boiled away and the vapor pressure within the barrel and the surface tension force kept the droplet of hydrogen
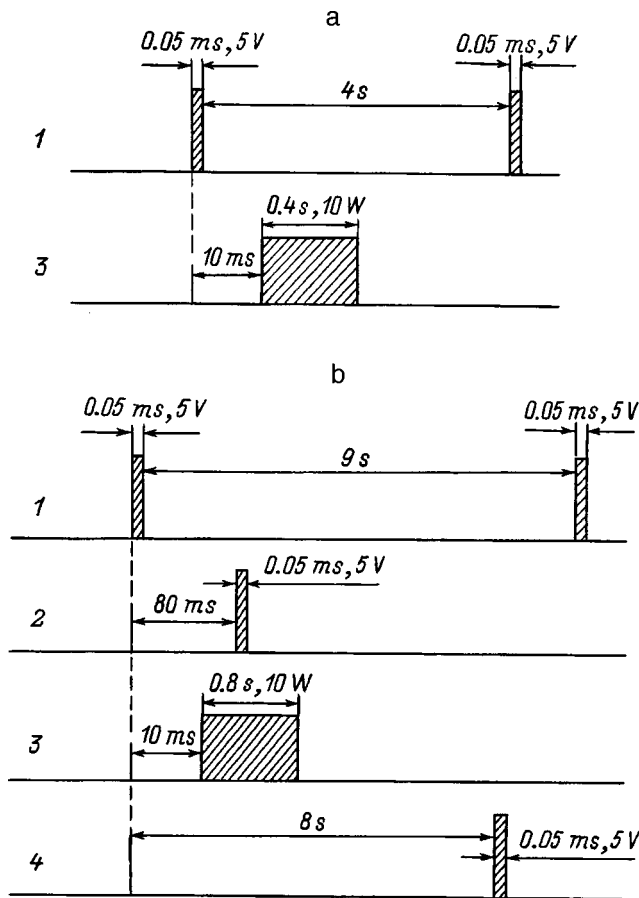
FIG. 3. Diagrams of the pulse sequence during forming and injection of pellets in the extrusion (a) and melting (b) modes: *1* — driver gas valve open, *2* — gate valve closed, *3* heater on, *4* — gate valve open.

FIG. 4. The temperature of the mold as a function of time during the extrusion (a) and melting (b) modes: *1* — pellet firing and heating of the mold, *2* — solidification of hydrogen, *3* — cooling of the mold and hydrogen pellet.

from spreading out along the walls. Over the 1–2 s after melting of the hydrogen and the filling of the barrel with the liquid, the hydrogen froze in the pores and no longer penetrated into the barrel. The porous structure served simultaneously as a thermal valve and liquid hydrogen dispenser. Then, over 6–8 s the hydrogen froze in the barrel and the gate valve was opened. 0.5 s after this a regular shot was fired. The driver gas and the gas formed by impact and vaporization of the pellet on the flange of the diagnostic chamber were pumped out and the cycle was repeated.

The temperature probe signals shown in Fig. 4 permitted an evaluation of the stability of pellet formation and the video record made it possible to estimate the pellet quality, dimensions, and velocity, as well as the reliability of injection. It is evident from Fig. 4 that the time to heat the mold in both modes was a fraction of a second, while the time to cool it and freeze the pellets in the barrel was different. In the extrusion mode the time to cool the mold to the initial temperature of 8 K was 5s during steady, periodic injection and this could be reduced to 3 s when the pellets were fired in short series of 3–5 pellets each. When pellets were fired with a separation of 3 s in longer series, the temperature of the mold rose, as shown in Fig. 4a, and the pellets broke up during acceleration. In the hydrogen melting mode, a delay of 3–5 s near the melting temperature (about 14 K) was established for cooling of the mold. This can be seen clearly
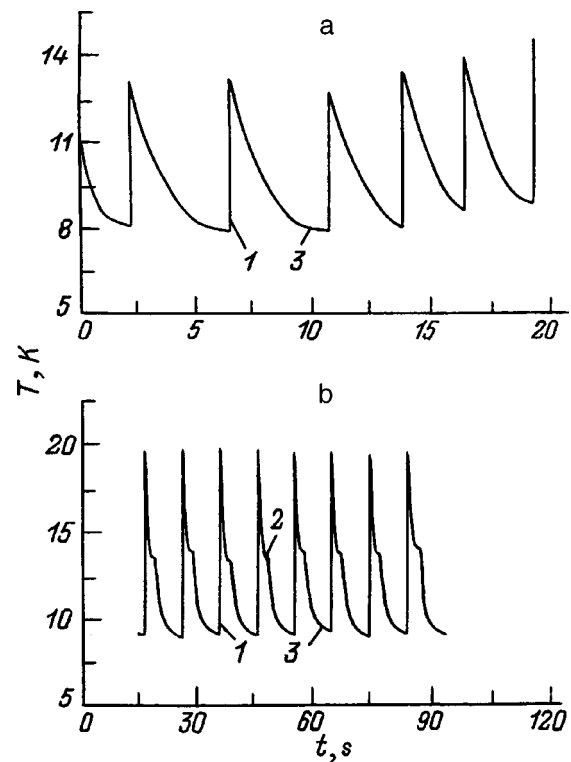
from the inflection of the curve in Fig. 4b and is related to a phase transition of the liquid hydrogen into the solid. Ultimately, in this mode the minimum time for cooling the mold to a temperature of 8 K after a shot was 8 s. In both modes helium was fed through the heat exchanger of the mold at a rate of about 12 liters/hour.

The pellets photographed during their flight had different forms: transparent if they were molded in the melting mode (Fig. 5a) and nontransparent if they were molded in the extrusion mode (Fig. 5b). Since both types of pellets were accelerated to 1–1.2 km/s by a gas at a pressure of up to 6 MPa without fracture and knocked small pieces out of the diagnostic chamber flange upon impact, we may assume that they were sufficiently durable to pass through a differential pumping system to a plasma without significant mass loss. A final confirmation of this will, however, will be provided by a suitable experiment.

More than a thousand pellets were formed and accelerated in the injector over a number of series with small time intervals between series. The main results are shown in Table I. The injection reliability was calculated as the ratio of unbroken pellets to the total number of pellets in a series. The velocities of the particles in a series were estimated from the position of a pellet with respect to a ruler located in the frame and from the time delay between the flash and the time the pellet passed by the laser beam at the detector. The velocities varied between 1 and 1.2 km/s for acceleration by helium at a pressure of 6 MPa. Their diameter was 2.1–2.4 mm and their length varied between 3 and 4 mm. The reduc-
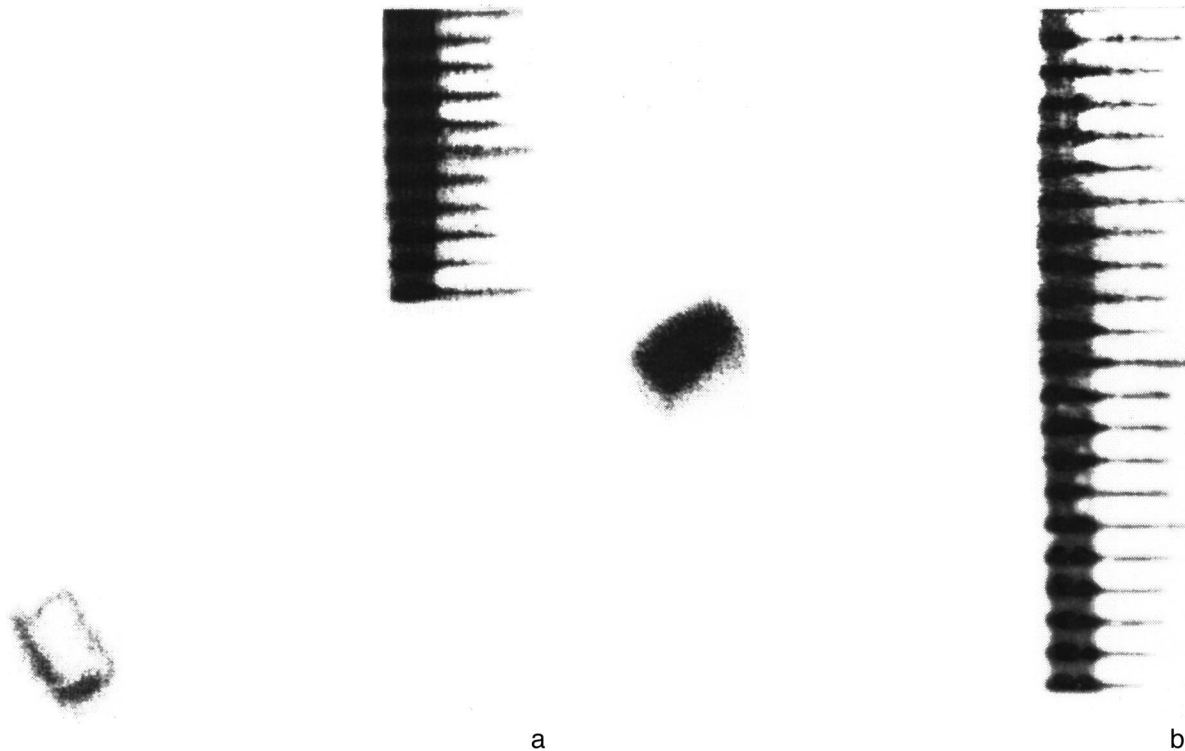
FIG. 5. Pellets formed in the extrusion (a) and melting (b) modes and accelerated to 1.2 km/s.

tion in the pellet diameter compared to the inner diameter of the barrel (2.7 mm) is explained by friction and sublimation from the side surface as the pellet moves in the barrel and is in good agreement with experiments on other injectors[9] and with model calculations.

CONCLUSION

The first stage of testing a ten-barrel injector for steady-state, periodic injection of pellets of thermonuclear fuel has been completed. The porous pellet mold has made it possible to inject solid hydrogen pellets at arate of 0.1–0.2 Hz from a single barrel. At any time there are only a few pellets in the injector, which makes it more attractive for tritium operation than extrusion injectors which contain hundreds of pellets. In addition, there are no moving parts in the injector except the valves, so it is more reliable and more easily repaired and can be regarded as an engineering device for steady-state thermonuclear reactors. After a mathematical model of the pellet formation is developed, the experiments will be continued with the aim of raising the injection rate to 0.5–1 Hz from a single barrel. Then it is planned to inject pellets from all 10 barrels in order to obtain a maximum firing rate.

TABLE I. Series of hydrogen pellets injected from a single barrel for different cycle parameters.

Injection cycle period, s,	Number of pellets in series	Hydrogen pressure in mold, MPa	Driver gas pressure, MPa	Average velocity, m/s	Reliability of injection, %
5	12	3	6	1100	92
6	17	3	6	1200	76
6	25	3	6	1000	88
7	22	3	6	1000	84
5−7	76	3	6	1060	85
8	15	0.1	6	1200	67
9	26	0.1	7	1200	62
9	55	0.1	6	1100	60
8−9	84	0.1	6	1100	62
10	22	0.1	5	700	64
11	26	0.1	6	1000	73
8−12	80	0.1	6	950	64
9−10	138	0.1	6	1050	60
9−14	211	0.1	6	1000	58

from the TOO ''Applied Physics'' for help in preparing the injector.

[1] S. Milora, W. Houlberg, L. Lengyel, and V. Mertens, Nucl. Fusion **35**, 657 (1995).
[2] B. Kuteev, A. Umov, I. Viniar *et al.*, Plasma Devices Op. No. 2, 193 (1994).
[3] S. Combs, Rev. Sci. Instrum. **64**, 1679 (1993).
[4] I. V. Vinyar, B. V. Kuteev, S. V. Skoblikov *et al.*, Zh. Tekh. Fiz. **65**, 167 (1995) [Tech. Phys. **40**, 723 (1995)].
[5] I. V. Vinyar, S. V. Skoblikov, and P. Yu. Koblents, Pis'ma Zh. Tekh. Fiz. **23**(20), 43 (1997) [Tech. Phys. Lett. **23**(10), 793 (1997)].
[6] S. Combs, S. Milora, C. Foster *et al.*, in *Fusion Technology, Proceedings of the 14th Symposium*, Pergamon, Oxford (1986), Vol. 1, pp. 355–360.
[7] I. V. Vinyar, Inventor's Certificate No. 1611139 [in Russian], publ. Byull. Izobret. **15**, 223 (1997).
[8] I. V. Viniar and S. Sudo, Rev. Sci. Instrum. **68**, 1444 (1997).
[9] S. Milora *et al.*, Nucl. Fusion **20**, 1491 (1980).

Translated by D. H. McNeill

Kinetics of an irreversible transition to a normal state by a multiconductor superconducting composite

V. R. Romanovskiĭ

Kurchatov Institute Russian Science Center, 123182 Moscow, Russia
(Submitted June 3, 1997; resubmitted January 26, 1998)
Zh. Tekh. Fiz. **69**, 40–51 (May 1999)

A model that takes into account the discrete character of a multiconductor superconducting medium is used to study the characteristic features of transition processes which occur when a region of normal conductivity appears. This analysis is based on numerical simulations of the propagation of a normal zone inside a superconducting region consisting of rectilinear superconductors, with a model for the kinetics of the irreversible loss of the superconducting properties of a single-ply, large radius winding. The thermophysical aspects of the irreversible propagation of a thermal perturbation within the transverse cross section of a multiconductor superconducting medium are analyzed. The major features of the phenomena taking place during the loss of superconductivity by a discrete region are formulated as functions of the conditions for heat transfer between its elements, the character of the perturbation, the magnitude of the flowing current, and the conditions for stabilization. The influence of longitudinal heat conduction on the magnitude of the transverse propagation velocity of a normal zone is examined. © *1999 American Institute of Physics.* [S1063-7842(99)00805-3]

INTRODUCTION

Keeping a superconducting magnetic system working after a section with normal conductivity develops within it and the subsequent propagation of this section along the winding is one of the major specifications for superconducting magnetic systems and their protective circuits. Protective measures acquire a special importance for the construction of superconducting magnetic systems with a high level of stored energy, which if released, can destroy the magnetic system. Thus, considerable attention continues to be paid to studies of the transition processes taking place in superconducting magnetic systems after a section with normal conductivity (a normal zone) appears in them.

Many theoretical results on transition processes have, to one degree or another, been obtained on the basis of models of a continuous medium. In this case, the superconducting magnetic system is regarded as an anisotropic continuum with averaged characteristics. As a result of this approximation, the spatial–temporal development of a transition process within the winding can be described by a model of an expanding ellipsoid[1–3] with its major axis directed along the axis of the wire. The following relationship holds between the velocities in the transverse V_y and longitudinal V_x directions with which the dimensions of this region increase: $V_y = V_x\sqrt{\lambda_y/\lambda_x}$, where λ_x and λ_y are the thermal conductivities in the corresponding directions. In this model, V_x is found by solving the corresponding one dimensional problem of the transition of a single superconducting composite wire to a normal state.[3] This approximation, however, does not provide a complete description of the features of the transition processes in a superconducting magnetic systems including its discreteness.[4,5] The characteristic mechanisms responsible for the propagation of a normal zone in regions with multiple winding structures are thereby ignored.

STATEMENT OF THE PROBLEM

We shall consider the propagation of a normal zone within a cooled superconducting medium with a discrete structure consisting of rectilinear superconducting composites (a superconductor in a normally conducting matrix), which is initiated by a powerful external thermal perturbation source. In order to simplify the analysis, we shall assume that the current in each element of the composite is constant, while their thermal and electrical properties are independent of the temperature and magnetic field. For this reason we shall neglect the temperature variation in the transverse cross section of each wire. We set the coordinate origin in the center of the segment with the initial temperature perturbation and describe the symmetric propagation of heat within adjoining composite superconductors, separated from one another by a finite thermal resistance, by a system of equations of the form

$$C\frac{\partial T_k}{\partial t} = \lambda\frac{\partial T_k^2}{\partial x^2} - \frac{hp}{S}(T_k - T_0) + \frac{I^2}{S^2}\rho(T_k)$$

$$-\begin{cases} \dfrac{P}{SR}(T_1 - T_2), & k=1 \\[2ex] \dfrac{P}{SR}(2T_k - T_{k-1} - T_{k+1}), & k=\overline{2, N-1}, \\[2ex] \dfrac{P}{SR}(T_N - T_{N-1}), & k=N \end{cases} \tag{1}$$

with the initial and boundary conditions

$$T_k(x,0)=\begin{cases} T_i, & 0<x<x_0, \quad k=k_i, \quad i=1,2,\dots, \\ T_0, & x_0\leq x\leq 1, \quad k=k_i, \\ T_0, & 0\leq x\leq 1, \quad k\neq k_i, \end{cases}$$

$$\frac{\partial T_k}{\partial x}(0,t)=0, \quad T_k(1,t)=T_0. \tag{2}$$

Here $k=1,\dots,N$ is the number of a wire in the composite; C is the specific heat per unit volume of the kth element; λ is its longitudinal thermal conductivity; h is the heat transfer coefficient; p is the cooled perimeter; S is the transverse cross sectional area; P is the contact perimeter between two neighboring wires; R is the thermal contact resistance; I is the transport current in each wire; T_0 is the temperature of the cooled medium; T_i is the initial temperature of the thermal perturbation, with size x_0; and, $\rho(T_i)$ is the effective resistivity of the superconducting composite including the existence of a region where the current is split up among segments in the superconducting and normal states in the kth wire,[2,3]

$$\rho(T_k)=\rho_0(T_k)$$

$$\times\begin{cases} 1, & T_k>T_{CB}, \\ \dfrac{T_k-T_C}{T_{CB}-T_C}, & T_C\leq T_k\leq T_{CB}, \\ 0, & T_k<T_C=T_{CB}-(T_{CB}-T_0)\dfrac{I}{I_C}, \end{cases}$$

where ρ_0 is the resistivity of the matrix, and I_C and T_{CB} are the critical parameters of the superconductor.

This system of equations describes the collective process by which a thermal perturbation with a given initial temperature and extent propagates within a multiconductor composite and causes the formation of a local region with normal conductivity within the k_ith element of the composite. The thermal interaction of the wires with one another is described by a model which assumes a linear steady temperature distribution within the intercontact region, which is true for a thin insulating spacer.[6] A finite difference method was used for determining the instantaneous temperature of all the elements of the composite and the corresponding propagation velocity of the segments with a normal conductivity.[7] Here we considered only those perturbations which necessarily cause an irreversible transition of the superconductor to a normal state (so-called supercritical perturbations[8]).

For formulating the basic physical behavior during the irreversible propagation of a normal zone within this composite, we reduce the overall number of initial parameters using the dimensionless variables $X=x/L_x$, $i=I/I_C$, $\tau=\lambda t/(CL_x^2)$, and $\Theta_k=(T_k-T_0)/(T_{CB}-T_0)$, where $L_x=[\lambda S^2(T_{CB}-T_b)/I_C^2\rho_0]^{1/2}$. Then Eqs. (1) and (2) take the form

$$\frac{\partial\Theta_k}{\partial\tau}=\frac{\partial^2\Theta_k}{\partial X^2}+i^2r(\Theta_k)-\frac{1}{\alpha}\Theta_k$$

$$-\begin{cases} \omega(\Theta_1-\Theta_2), & k=1, \\ \omega(2\Theta_k-\Theta_{k-1}-\Theta_{k+1}), & k=\overline{2,N-1}, \\ \omega(\Theta_N-\Theta_{N-1}), & k=N, \end{cases}$$

$$\Theta_k(X,0)=\begin{cases} \Theta_i, & 0<X<X_0, \quad k=k_i, \\ 0, & X_0\leq X\leq L, \quad k=k_i, \\ 0, & 0\leq X\leq L, \quad k\neq k_i, \end{cases}$$

$$\frac{\partial\Theta_k}{\partial X}(0,\tau)=0, \quad \Theta_k(L,\tau)=0.$$

Here

$$r(\Theta_k)=\begin{cases} 1, & \Theta_k>1, \\ \dfrac{\Theta_k-1+i}{i}, & 1-i\leq\Theta_k\leq1, \\ 0, & \Theta_k<1-i, \end{cases}$$

and α and ω are dimensionless parameters accounting for heat transfer into the cooling agent and the thermal coupling among the wires, respectively.

α (the stabilization parameter) is calculated using the standard expression[2,3] $\alpha=I_C^2\rho_0/hpS(T_{CB}-T_0)$, while the dimensionless thermal resistance ω is calculated in conventional dimensionless form as $\omega=PS(T_{CB}-T_0)/(RI_C^2\rho_0)$. In particular, for $\omega=0$ this system of equations takes a form that describes the change in the thermal state of a single superconducting composite.[8]

This simplified description of the transition of a superconducting magnetic system to a normal state allows us, within the generally accepted terms of the theory of thermal stabilization, to carry out a generalized analysis of the major features of the thermal processes which take place during loss of the superconducting properties of a current carrying element of a superconducting magnetic system. Below we discuss the results of some numerical simulations which reflect the qualitative behavior of the transition of a superconducting discrete composite to a normal state. In doing these calculations, in order to reduce edge effects owing to specifying a boundary condition of the first kind at $X=L$, the extent of the computational region was taken to be $l=500$. As the calculations show, this corresponds to an essentially infinite longitudinal extent of the computational region. Without loss of generality, in most cases it was assumed that at the initial time a normal zone appears in the first wire ($k_i=1$) owing to powerful local heating at a dimensionless level of $\Theta_i=10$.

FORMATION OF A KERNEL WITH NORMAL CONDUCTIVITY INSIDE A SUPERCONDUCTING MEDIUM WITH A DISCRETE STRUCTURE

Figure 1 illustrates the propagation within a multiconductor composite of a heat releasing region whose boundary is described by an equation $\Theta_k(X_{k,n},\tau)=1$ corresponding to the phase interface between the superconducting and nonsuperconducting states. The initial parameters were specified to be $\alpha=100$, $i=0.5$, $\omega=0.1$, and $X_0=10$. Figure 2 shows plots of the instantaneous velocities $dX_{k,n}/d\tau$ of the isotherms
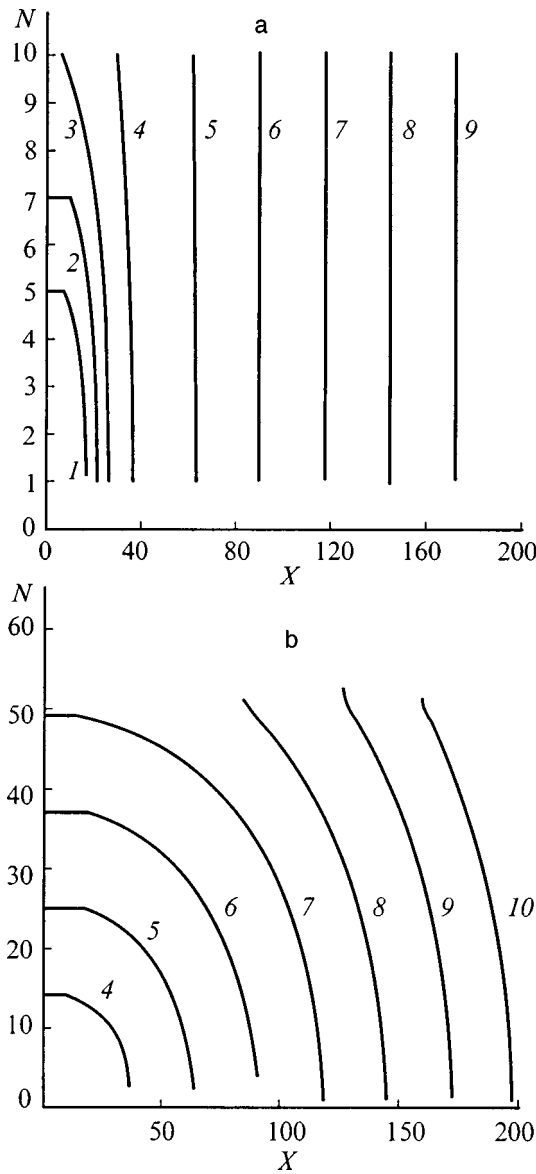
FIG. 1. Penetration of a normal zone inside a superconducting discrete composite with different numbers of wires following local development of this zone in element number $k = 1$. $N = 10$ (a), 51 (b). $\tau = 10$(*1*), 20 (*2*), 30 (*3*), 50 (*4*), 100 (*5*), 150 (*6*), 200 (*7*), 250 (*8*), 300 (*9*), and 350 (*10*).

$\Theta_k(X_{k,n}, \tau) = 1$ with which it moves along wire number $k = 1$, where the normal zone was initiated. Calculations were done for $\omega = 0.2$ for two characteristic values of the stability parameter corresponding to ''poor'' ($\alpha = 100$) and ''good'' ($\alpha = 2$) heating for different values of the current and total number of wires. For the case $\alpha = 100$, Fig. 2 also shows the propagation velocity of theisotherm $\Theta_k(X_{k,n}, \tau) = 1$ along elements of a composite with $N = 51$.

These results demonstrate clearly the features of the transition process and the formation of a normally conducting kernel inside a superconducting medium with a discrete structure. First of all, one can see that there are two characteristic regimes for the propagation of a normal zone along a multiconductor composite. In the initial stage of the formation of a resistive region, the velocity of the normal zone is essentially nonstationary. Its duration depends on the cooling conditions, current load, the thermal coupling among the

wires, and the number of wires. In the second stage, the velocities of the normal zones in each wire successively stabilize and reach constant values equal to the corresponding velocity of a normal zone in a single composite. This corresponds to so-called quasistationary states, when the thermal instability propagates along a composite superconductor in the form of a characteristic thermal wave.[2,3] Here the approach to a stationary value in each element of the composite is asymptotic. This state is reached primarily in the wire where the perturbation was initiated. The propagation of the thermal perturbation in the other wires is not only more nonstationary in character, but may take place at higher velocities than in an analogous single wire. Entirely trivially, this happens because a new normally conducting segment develops in wires with an elevated background temperature. Thus, in the first wire the velocity of the normal zone approaches its asymptotic value from below. At the same time, in the other wires this process has a different type of convergence, both from below and from above. In terms of the model of an expanding ellipsoid, this description of the kinetics of the normal zone cannot hold, since the velocity of any point on an ellipse lying in an upper ''layer''of it is always higher than the velocity in a lower ''layer.''

It should also be noted that the increase in the size of a resistive region in a superconducting medium with a discrete structure has a number of other features. First of all, the boundary of the region in a normal state has the shape of a truncated ellipse. This is because of the finite transition time for each wire to a normal state. Obviously, the worse the thermal coupling between the windings is, the more cut-off the shape of the resistive region will be. Second, the distinctive feature of the transition to a stationary velocity of the normal zone is the local transition of all the wires to a normal state. Thus, at first there is a gradual flattening of the boundary separating the superconducting and nonsuperconducting regions, which then transforms into a straight line. And finally, we note the most important limitation on the validity of the continuous medium model. As the calculations shown in Figs. 3 and 4 ($N = 20$) show, at currents close to the so-called stationary stabilization current (in the terminology used here it equals $i_s = (\sqrt{1 + 8\alpha} - 1)/2\alpha$) in a superconducting medium with a discrete structure, it turns out that states are possible for which an irreversible transition of the entire composite to a normal state does not take place. For the case $\alpha = 100$, $i = 0.5$, $\omega = 0.1$, and $X_0 = 10$, Fig. 3 shows instantaneous propagation velocities of the isotherm $\Theta_k(X_{k,n}, \tau) = 1$ at which it moves along the wires for currents ranging from $i = 0.14$ to 0.2 that only slightly exceeded the stationary stabilization current ($i_s = 0.1356$). Here also the dashed curves show the corresponding velocities of the isotherm $\Theta_k(X_n, \tau) = 1$ for a single composite and the inset to Fig. 3 illustrates the kinetics of the penetration of the normal zone in a transverse cross section perpendicular to the main propagation direction of the normal zone. It is quite clear that for $i \geqslant 0.18$ the above features of an irreversible transition of the entire superconducting composite to a normal state are fully valid. At the same time, for currents $i < 0.18$ situations may occur in which a limited number of wires inside a composite enter the normal state. Thus, for
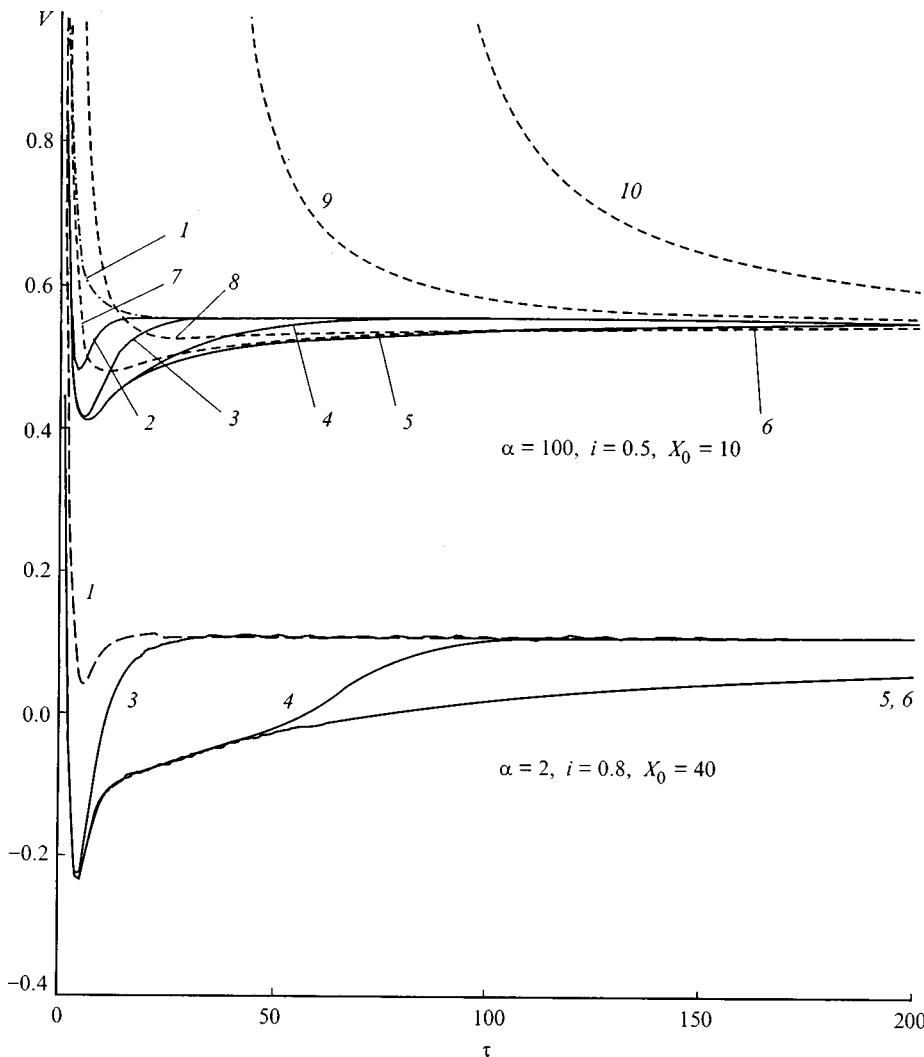
FIG. 2. Kinetics of a normal zone over the elements of a composite: $k=1$, $N=1$ (*1*), 2 (*2*), 3 (*3*), 5 (*4*), 10(*5*), 51 (*6*); $N=51$, $k=2$ (*7*), 3 (*8*), 10 (*9*), 20 (*10*).

these parameters, when $i=0.17$ a resistive region exists in the steady state only in three wires, and for $i=0.16$ and 0.15, in two. When the current is reduced further, quasistationary states may not develop. The existence of such states means that the propagation velocity of the normal zone inside a discrete composite is lower than the corresponding velocity in a single composite, and this difference becomes greater as the current is lowered.

The existence of these states also depends on the character of the thermal coupling among the elements of the composite. Figure 4 shows curves illustrating the kinetics of a normal zone for a composite with different values of ω. It is clear that for $\alpha=100$ and $i=0.5$, ω only affects the time of the transition process. However, for $\alpha=2$ and $i=0.8$ ($i_s = 0.7808$), the conditions for the development and propagation of a normal zone inside a discrete superconducting medium may undergo a change. Thus, for $\omega=0.01$, only one wire goes into a normal state, for $\omega=0.1$, two, and for $\omega=0.2$, the entire composite as a whole.

PROPAGATION OF A THERMAL PERTURBATION IN A TRANSVERSE CROSS SECTION OF A MULTICONDUCTOR SUPERCONDUCTING COMPOSITE

Figures 5–7 show numerical simulations of the development and propagation of segments with normal conductivity within the transverse cross section of a superconducting discrete medium for $\alpha=100$. The calculations were done for different currents, temperatures, and sizes of the initial perturbation, heat transfer conditions among the elements of the composite, and numbers of elements. The boundary of the heat release region was determined by solving the equation $\Theta_k(X_{k,n},\tau)=1$. Then, from the condition $\Theta_k(0,\tau_{q,k})=1$ it is easy to find the time for the normal zone to appear in the kth wire. These temporal states are indicated in Figs. 5–7 by symbols that are joined together. This makes it possible to approximate the discrete process of forming new, normally conducting segments with the corresponding continuous equations.

The curves in Fig. 5 describe the kinetics of the transition process in composites with different numbers of elements for three dimensionless currents with $X_0=50$ and $\omega=0.1$. The temperature of the initial perturbation for $i=0.4$ and 0.8 was specified to be $\Theta_i=10$, while for $i=0.3$ the transition process was calculated for several values of Θ_i. Figure 6 shows plots of $N(\tau)$ when the size of the initial perturbed segment is varied for $i=0.3$, $N=10$ and 20, and $\Theta_i=10$. These results clearly demonstrate the possible ranges of variation in the initial parameters for which the superconducting properties of the entire composite are lost
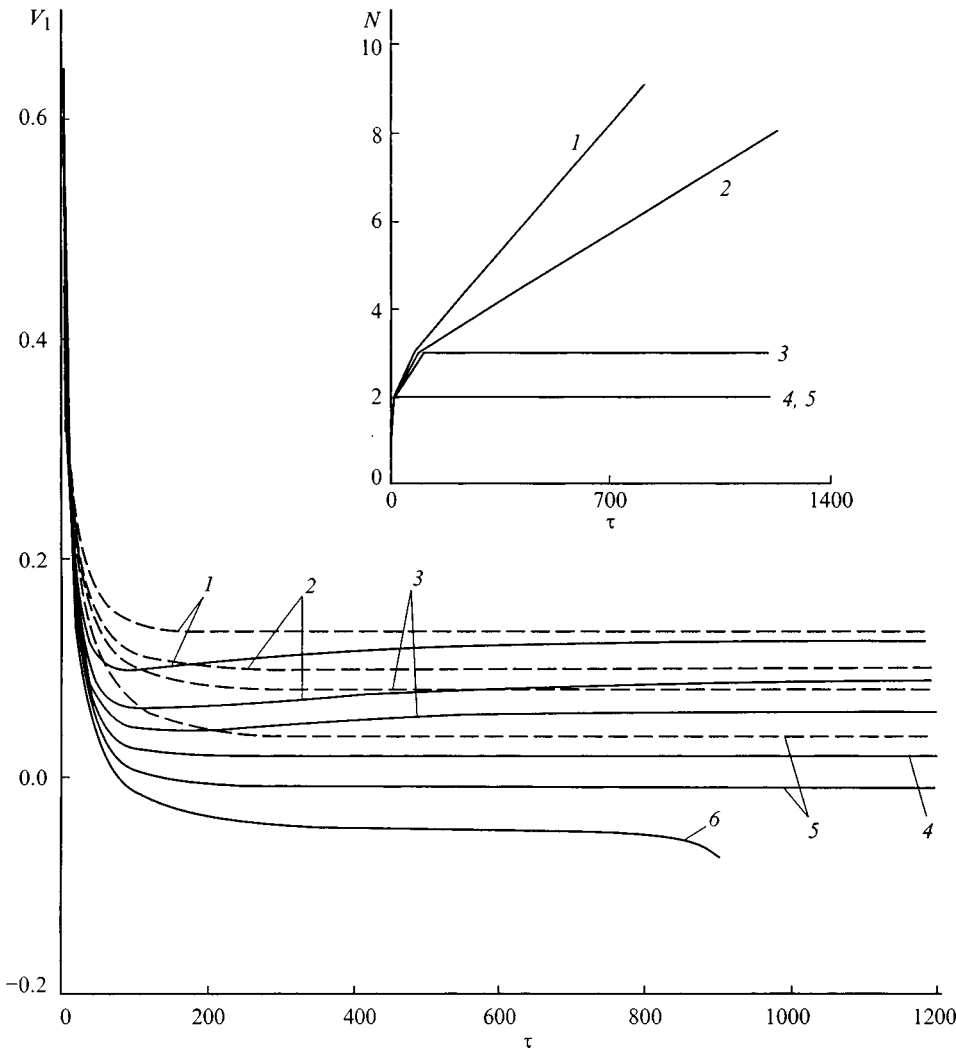
FIG. 3. Time dependence of the propagation velocity of a normal zone in wire number $k=1$ near the stationary stabilization current: $i=0.2(1)$, 0.18 (2), 0.17 (3), 0.16 (4), 0.15 (5), 0.14 (6).

irreversibly, and within which the major features of this process will take place. In particular, it should be noted that the evolution of the transition process in the transverse cross section of the composite ceases to depend on the size of the normal zone in the longitudinal direction, if that dimension exceeds 50 (dimensionless units) at the initial time. Here the power of the external perturbation essentially has no effect on the kinetics of the transition process. In addition, except for edge effects associated with specifying a finite number of wires in the composite, it is desirable that their total number be at least 10. Thus, the numerical simulations presented below have been done for $N=20$, $X_0=50$, and $\Theta_i=10$.

Figure 7 shows plots of $N(\tau)$ for composites with different thermal coupling parameters among the wires as the current flowing in them is varied over a wide range. The dashed curve shows calculations for the corresponding case of the simultaneous appearance of a localized segment with normal conductivity in several wires.

By comparing these calculations, it is easy to establish the basic behavior which controls the characteristic features of the irreversible propagation of a thermal perturbation in the transverse cross section of a superconducting multiconductor medium.

First, during the formation of the phase boundary there are periods of time when the growth in the perpendicular dimensions of the resistive region proceeds at unsteady velocities. These occur in the initial and final stages of the time variation $N(\tau)$. As a rule, they are quite brief. But they last longer for lower current and power, and for smaller sizes of the external perturbation, and when the conditions for heat transfer among the elements of the composite are poorer. In the intermediate range of times for the transition process, the velocity at which the dimensions of the resistive region increase is constant, at a value which does not depend on the character of the initial perturbation (more precisely, because of the computational model used here, it approaches this value asymptotically). Thus, in this case we can speak of the development of several quasistationary states, for which the conditions of existence will depend only on the intrinsic properties of the elements of the composite.

Second, the end result of the penetration of a normal zone into a multiconductor medium turns out to depend on the conditions for dissipation of the Joule heating. Thus, the curves in Fig. 5 correspond to states where all the superconducting region must go into a normal state owing to the interaction of the perturbation in only the first element. On the other hand, Fig. 7 shows that when the instability is initiated in a single wire at low currents (in a more rigorous
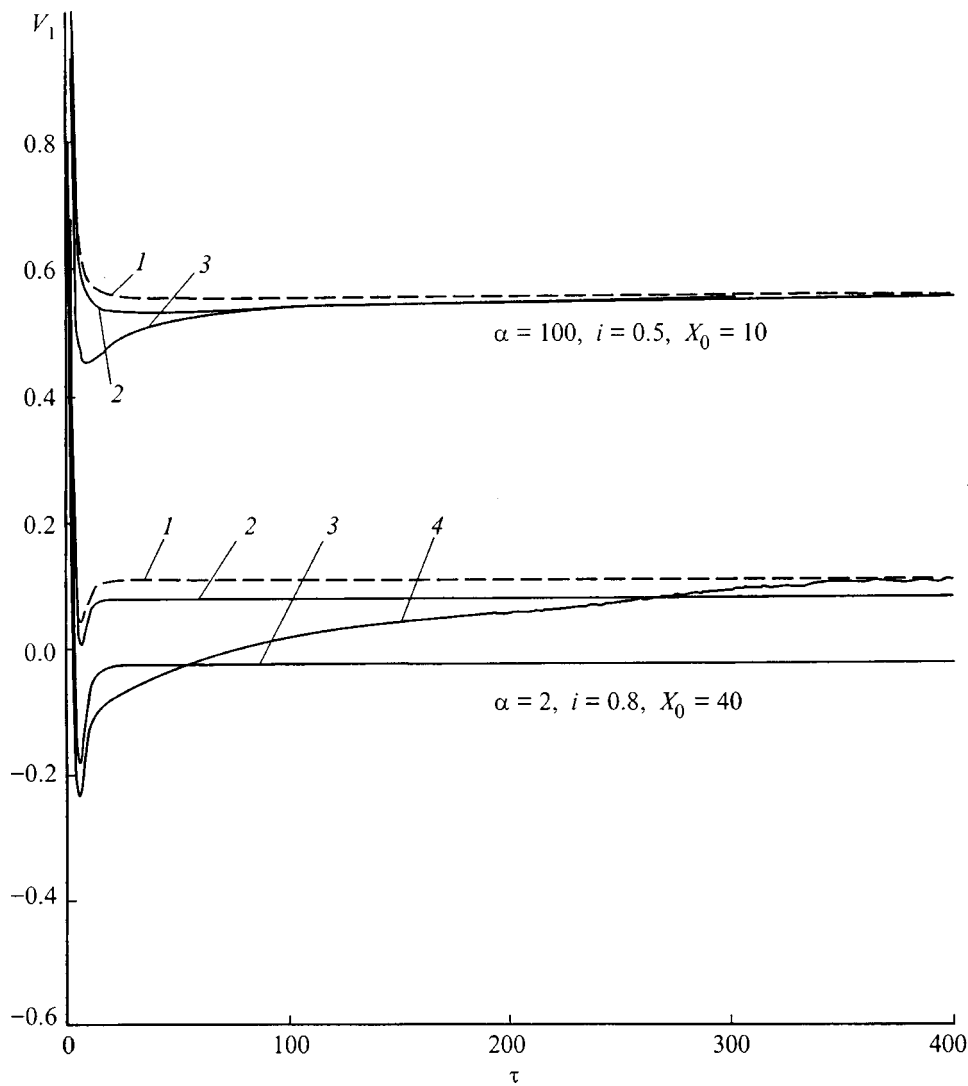
FIG. 4. Time variation in the propagation velocity of a normal zone in wire number $k=1$ for different heat transfer conditions among the wires; $\omega=0$ (*1*), 0.01 (*2*), 0.1 (*3*), 0.2 (*4*).

formulation, near the stationary stabilization current i_s), irreversible propagation of the thermal perturbation over the entire cross section of the composite may not occur. Thus, for the given parameters, states of this sort occur for $i < 0.16$ ($i_s = 0.1356$). In this case, complete loss of the superconducting properties of the composite can take place when several wires undergo a simultaneous transition to a normal state. The dashed curve in Fig. 7 shows the corresponding calculations for $i = 0.15$ when a normal zone develops at the initial time in five (curve *5*) or fifteen (curve *6*) wires. It is easy to note that for the given initial conditions there is both a substantial increase in the time of the unsteady regime, and the possibility that the steady states don't exist at all. This must be taken into account in the corresponding experiments.

The results discussed above allow us to formulate a formal rule for determining the transverse propagation velocity of a thermal perturbation in a superconducting medium with a discrete structure. In terms of the dimensionless units of this paper, it is given by $V_y = 1/(\tau_{q,k} - \tau_{q,k-1})$. Here $\tau_{q,k}$ and $\tau_{q,k-1}$ are the times for a transition to a normal state of the kth and $(k-1)$-st elements of the structure. The steady-state value must be determined after the initial nonstationary period has passed. As the curves in Figs. 5–7 show, this value

can be approximated with good accuracy by solving the problem of the transition to a normal state by a multiconductor composite with a finite number of elements.

In Fig. 8 the solid curves represent the calculated transverse propagation velocity of the normal zone as a function of current. The calculation was done for two characteristic values of the stability parameter: $\alpha=2$ for effective thermal stabilization (e.g., when intense cooling is present) and $\alpha=100$ for a low degree of thermal stabilization (e.g., when the cooling conditions are close to adiabatic). The dimensionless parameter for the thermal coupling among the wires was taken to be equal to 0.01 and 0.1. (For $\omega=0$ the trivial $V_y=0$ holds.) The plotted curves were calculated assuming that an instability at the initial time appears only in the first wire. Also shown here for clarity, as the dashed curves, are the corresponding variations in the longitudinal propagation velocity of a normal zone in a single composite, when the irreversible loss of its superconducting properties takes place against the background of a thermal wave moving at a constant velocity. Formally, this corresponds to the case $\omega=0$.

Figure 8 implies that a deterioration in the thermal coupling among the wires not only lowers the transverse propagation velocity of a normal zone, but also reduces the range
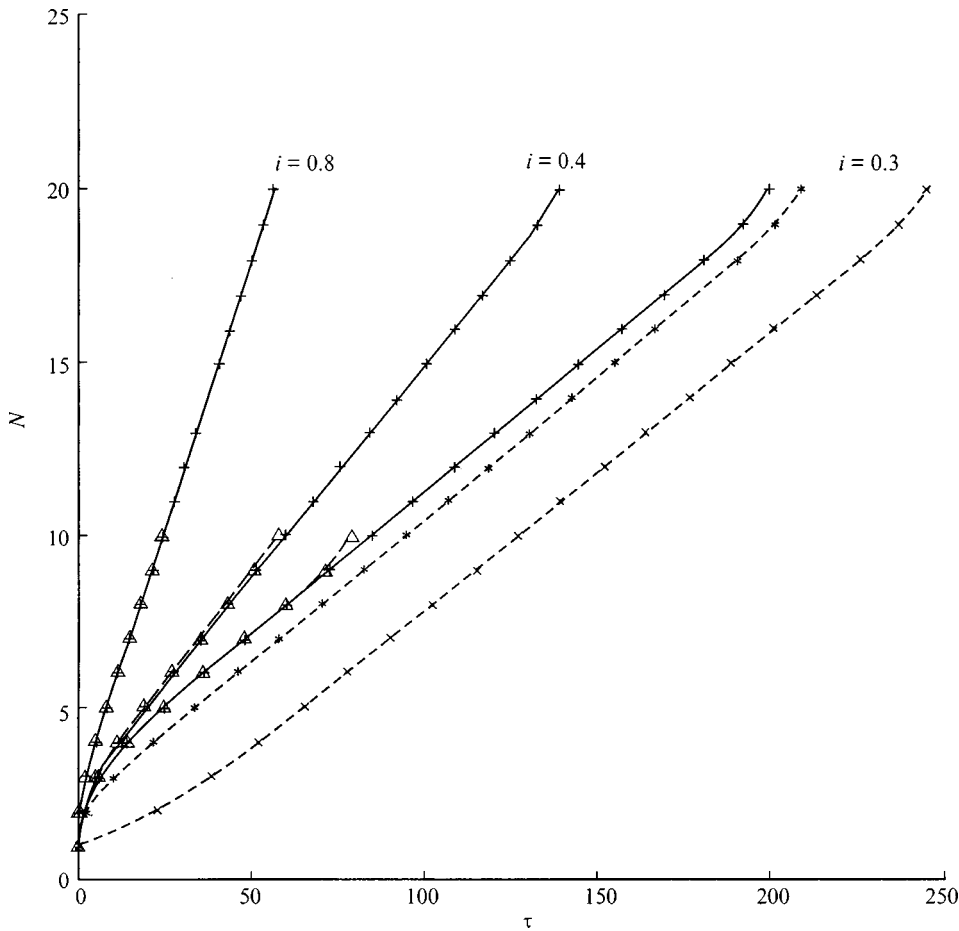
FIG. 5. Effect of the total number of wires in a superconducting composite and of the temperature of the initial perturbation on the kinetics of the irreversible loss of superconductivity in the transverse cross section: $N=10$, $\Theta_i=10$ $(---\Delta)$; $N=20$, $\Theta_i=1.0$ $(-+-)$, 5 $(---*)$, 1 $(---\times)$.

of currents over which a complete transition of a multiconductor composite to a normal state initiated by an instability in a single wire can take place. This effect is observed to the greatest extent in wires which are ''well'' stabilized from the standpoint of thermal stability. Besides this, for these stabilization regimes, the $V_y(i)$ curve can be approximated linearly as the current increases over a wide range of its variation. It becomes nonlinear in the neighborhood of the stationary stabilization current, near the critical current, and when ω is increased.

It follows from the above remarks that for a discrete superconducting region the formal ratio of the steady state velocity of the normal zone in the transverse cross section to its velocity in the longitudinal direction is not constant, as implied by expanding ellipsoid model. In this model,[3] the propagation velocity of the normal zone in a single thin superconducting composite,

$$V_x = \frac{J}{C}\left(\frac{\lambda_x \rho_0}{T_{CB}-T_0}\right)^{1/2}$$

is taken to be equal to the speed at which the major axis of the ellipsoid increases (J is the current density and C the specific heat per unit volume). Assuming in the simplest case that in a transition from a discrete medium to a continuum, all the parameters except the thermal conductivity are unchanged, it is easy to write down an expression for the velocity of the normal zone in the direction of the minor axis of the ellipsoid,

$$V_y = \frac{J}{C}\left(\frac{\lambda_y \rho_0}{T_{CB}-T_0}\right)^{1/2},$$

where λ_y is the average of the thermal conductivity over the transverse cross section.

With such a transition it is also possible to account for the change in the other properties of the composite. But this does not change the relationship among the velocities qualitatively, when $V_y/V_x=$ const over the entire range of variation in the current. As a result, this ratio is independent of the magnitude of the flowing current. At the same time, a direct numerical simulation yields a nonlinearity in the $V_y(i)$ curves, owing, first of all, to the presence of a special range of currents for which the conditions for penetration of the instability inside a multiconductor composite are not a consequence of the reason that appears evident at first glance and is associated with the trivial heating initiated by the perturbation in a single wire of the composite. Figure 9 shows the calculated ratio V_y/V_x for $\alpha=100$ as a function of the current in the case where a normal zone develops in the first wire. It can be seen that as the current rises, the nonlinear character of V_y/V_x is fully determined by a corresponding variation in the transverse propagation velocity for the normal zone. Thus, at first (at low currents) the increase in V_y/V_x has a noticeable tendency to rise and then its variation is not so strong. An improvement in the heat transfer between the elements of the composite is accompanied by a sharper change in the initial segment of these curves. How-
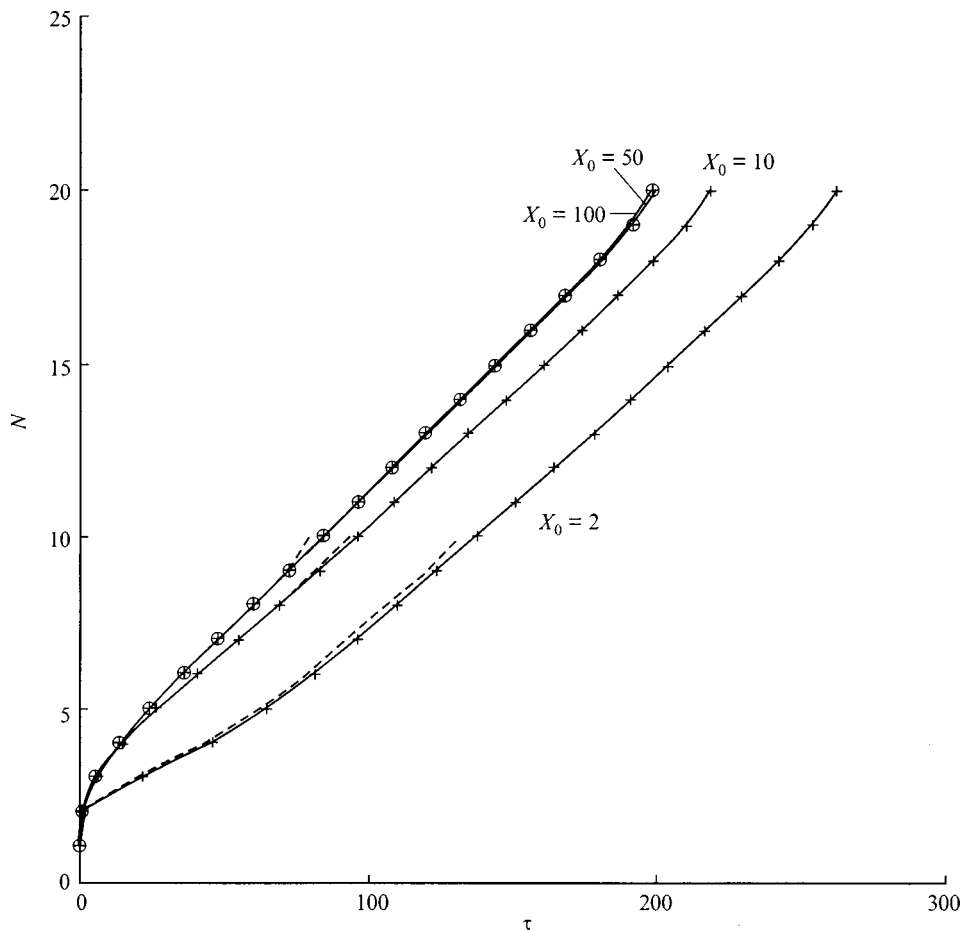
FIG. 6. Effect of the size of the initial perturbation on the penetration of a normal zone inside a multiconductor composite: $N = 10$ $(---)$; $N = 20$, $X_0 = 2$, 5, 50 $(-+-)$, 100 $(-\bigcirc-)$.

ever, in this case the range of currents where V_y/V_x has a sharp rise becomes smaller. Evidently, in the limit $\omega \to \infty$, for example, when there are no insulating spacers this range of currents will be absent in the $V_y/V_x(i)$ curves.

Besides these features which the expanding ellipsoid model ignores, that model also yields a variation for the transverse velocity of propagation of the normal region with the thermal conductivity of a single element of the composite in its longitudinal direction. This relationship is evident for a continuum, whose physical characters are not discontinuous, when a heat flux inside it propagates continuously in all directions. In the case of a superconducting region with discrete physical characteristics, the propagation of a normal zone takes place, first of all, in a geometrically specified direction resulting from the design features of the composite. In this case, because of the rapidity of the propagation of a normal zone along the wire, the time interval between two successive times of formation of new normally conducting segments depends, first of all, on the thermal properties of the insulating spacer and on the temperature of the wire in its hottest portion. Its magnitude is determined mainly by the Joule heating power and the conditions under which it is dissipated in the cooling agent and in the adjacent layers. Thus, longitudinal heat conduction in each of the elements of the composite should have little effect on the conditions for the appearance and transverse propagation of a normal zone. As an illustration of this, Fig. 10 shows the results of a numerical simulation for $\alpha = 100$ of the propagation of a per-

turbation in composites with thermal conductivities which differ by two orders of magnitude. It is based on solving Eqs. (1) and (2) with the term $\partial^2\Theta_k/\partial X^2$ replaced by a more general expression $\Lambda\partial^2\Theta_k/\partial X^2$, where Λ is the dimensionless thermal conductivity of the kth element. In solving the model problem, the stability parameter was set to $\alpha = 100$ and the current and the heat transfer conditions among the wires were varied. These calculations demonstrate clearly the extremely small contribution of longitudinal heat transfer to the kinetics of the penetration of a perturbation in the transverse cross section of a multiconductor composite. In particular, it has essentially no effect on the transverse velocity of a normal zone if the instability develops near the critical current. At lower currents its contribution become somewhat noticeable if there is poor thermal coupling among the wires.

CONCLUSIONS

This study has shown that the irreversible propagation of a normal zone inside a discrete superconducting medium has a number of features which are not described by the expanding ellipsoid model. These are based on the following behavior.

1. The resistive region formed as a result of the irreversible to a normal state of the superconducting elements of a composite consisting of superconductors separated by a finite thermal resistance has the shape of a truncated ellipse.
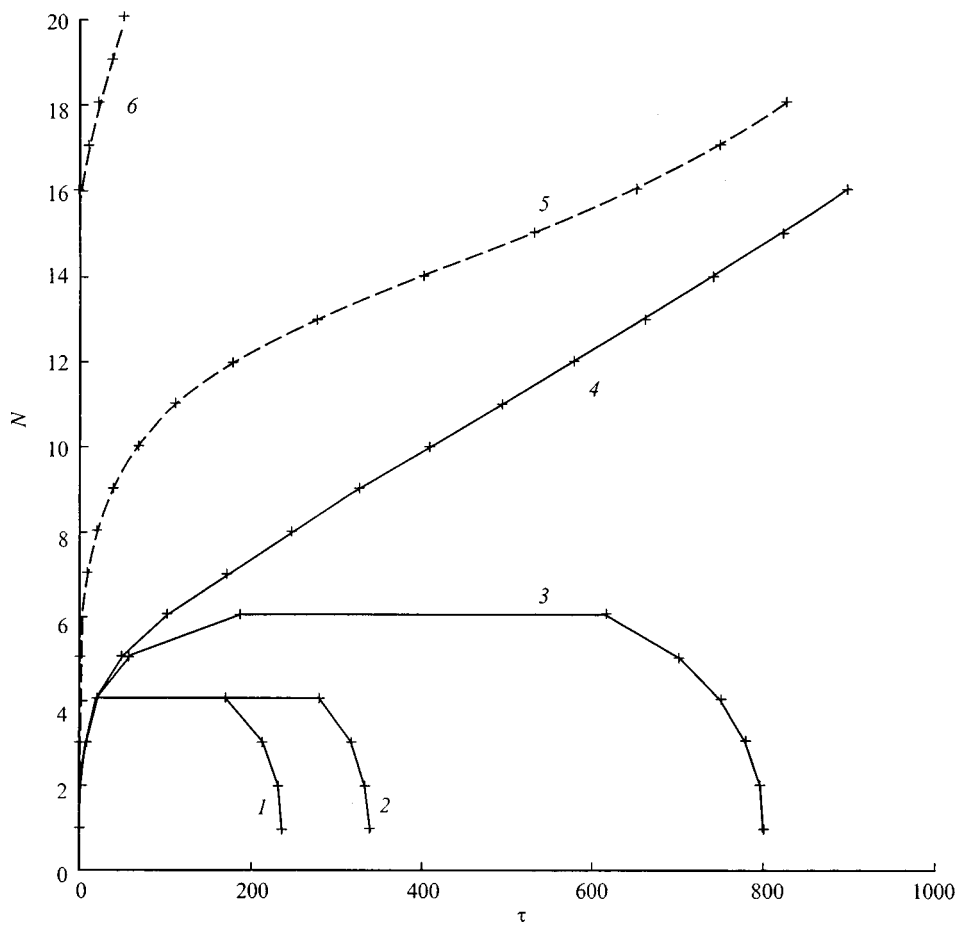
FIG. 7. Kinetics of the transition process for $\omega = 0.1$ near the stationary stabilization point: $i = 0.13$ (*1*), 0.14 (*2*), 0.15 (*3–6*), 0.16 (*4*).

2. The rate of growth of its principal axis (the propagation velocity of the normal zone in the longitudinal direction) is less than the corresponding value of the propagation velocity of a normal zone in a single composite. The latter is an asymptotic limit for all velocities at which a normal zone will propagate in each element of the composite. This limit is attained only after local transition of all the elements of the

composite to a normal state. Then the internal boundary separating the superconducting region from the nonsuperconducting region acquires a flat shape.

3. At currents close to the stationary stabilization currents, states may occur in which the superconducting properties are not lost in all the elements of the composite. As a result, a limited number of subregions with normal conduc-
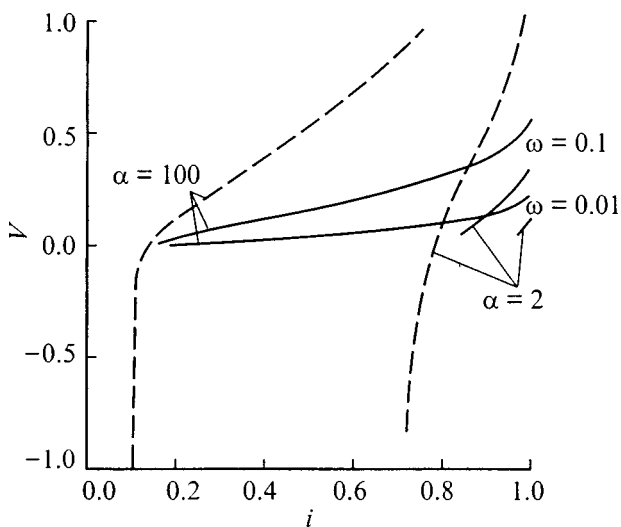
FIG. 8. The steady propagation velocities of a normal zone as functions of current: solid curves V_y, dashed curves V_x.
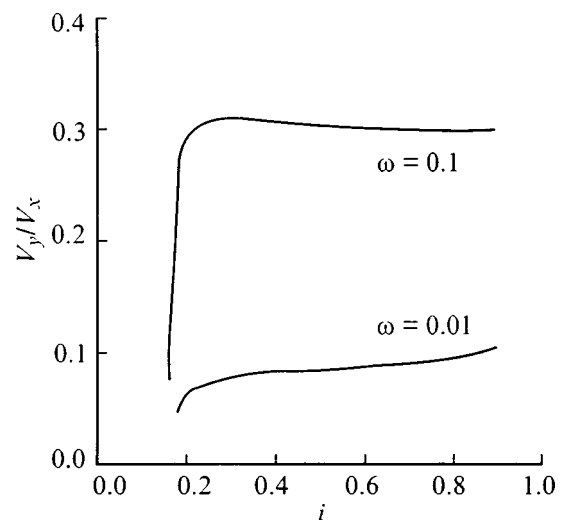
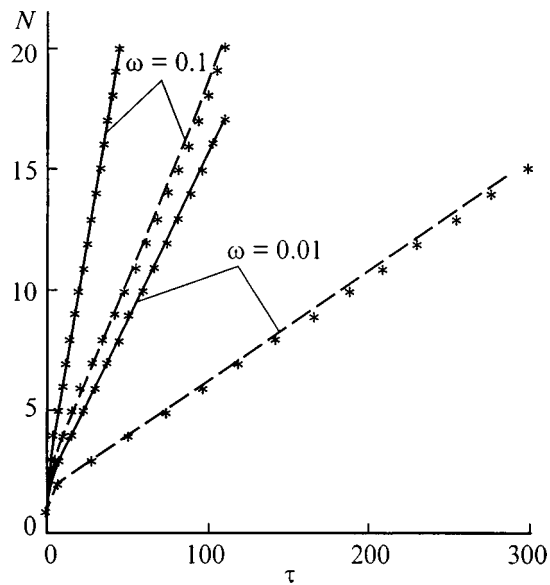FIG. 9. The ratio of the transverse and longitudinal velocities of a normal zone as a function of current.

FIG. 10. The effect of longitudinal thermal conductivity on the kinetics of the transition process: solid curves $i=0.9$, dashed curves $i=0.5$; $\Lambda=0.1$ (curves), 10 (stars).

tivity, within which a normal zone propagates at a constant velocity, can exist stably inside a superconducting medium with a discrete structure.

4. Near the stationary stabilization current, the way the transition process evolves depends to a substantial degree on the conditions of heat transfer among the elements of the composite. Thus, when a normal zone appears in a single element of a multiconductor medium with insulating spacers that have relatively high thermal resistances, the whole composite may not undergo a transition to the normal state. This effect is observed to the greatest extent in ''well'' stabilized conductors.

5. The propagation velocity of a normal zone in the transverse cross section of a multiconductor composite is essentially independent of the longitudinal thermal conductivity of its elements. It is determined primarily by the thermal properties of the insulator.

[1] Z. J. J. Stekly, Adv. Cryog. Eng. **8**, 585 (1963).
[2] V. A. Al'tov, V. B. Zenkevich, M. G. Kremlev, and V. V. Sychev, *Stabilization of Superconducting Magnetic Systems* [in Russian], Énergoatomizdat, Moscow (1984), 312 pp.
[3] M. Wilson, *Superconducting Magnets* [Oxford University Press, London (1983); Mir, Moscow (1985), 407 pp.].
[4] A. I. Rusinov, IEEE Trans. Magn. **MAG-30**, 2681 (1994).
[5] A. V. Gavrilin, IEEE Trans. Appl. Supercond. **AS-3**, 293 (1993).
[6] V. S. Zarubin, *Engineering Methods for Solving Heat Conduction Problems* [in Russian], Énergoatomizdat, Moscow (1983), 326 pp.
[7] V. M. Paskonov, V. I. Polezhaev, and L. A. Chudov, *Numerical Simulation of Heat- and Mass-Transfer Processes* [in Russian], Nauka, Moscow (1984), 286 pp.
[8] V. R. Romanovskiĭ, Izv. AN SSSR. Énergetika i Transport, No. 4, 115–119 (1984).

Translated by D. H. McNeill

Microwave breakdown initiated in a high-temperature superconducting film by thermal disturbances and defects

N. A. Buznikov and A. A. Pukhov

Joint Institute for High Temperatures, Scientific Center for Applied Problems of Electrodynamics, Russian Academy of Sciences, 127412 Moscow, Russia
(Submitted December 15, 1997)
Zh. Tekh. Fiz. **69**, 52–59 (May 1999)

The effect of thermal disturbances and nonsuperconducting defects on the microwave breakdown of a high-temperature superconducting film is investigated theoretically. The dependence of the critical energy of thermal disturbances on the surface microwave field is obtained. It is shown that in a wide range of values of the surface microwave field the critical energy of local disturbances is less than the critical energy of spatially extended disturbances. It is established that microwave breakdown on defects may be preceded by a stage of formation of a finite-size normal-phase region localized at the defect. The effect of the form and absorption coefficient of the defect on scenarios leading to the destruction of superconductivity and on the microwave breakdown field of a film at a defect are investigated. © *1999 American Institute of Physics.* [S1063-7842(99)00905-8]

INTRODUCTION

Microwave breakdown (destruction of superconductivity by microwave radiation) of high-temperature superconducting (HTSC) films is observed at quite high surface microwave fields $B \sim 10$ mT,[1–8] corresponding to the incident radiation power densities $P \sim 10^6$ W·cm^{-2}. The thermal mechanism of breakdown discussed for low-temperature superconductors[9,10] is associated with a sharp increase of the absorption coefficient of the film above the critical temperature T_c,[11,12] which results in the existence of a threshold breakdown power density P_b.[13] The spatiotemporal pattern of the destruction of superconductivity in a film at $P > P_b$ is related with the possibility of unbounded expansion of the normal-phase region, which occurs by propagation of a temperature autowave of S–N switching along the film.[13–16] A normal-phase region sufficient for the appearance of a temperature autowave can be initiated in a film by a thermal disturbance[13,17] with sufficiently high energy, exceeding a critical value E_c. The critical energy E_c is a complicated function of the properties of the HTSC film, the cooling conditions, and the duration and spatial extent of the disturbance, and in general it can be found only numerically. It will be shown below that simple analytic expressions for the critical energy can be obtained in the limiting cases of local and spatially extended disturbances. Another scenario leading to the formation of a normal-phase regions is related with the existence of normal (nonsuperconducting) defects in an HTSC film.[18,19] The destruction of superconductivity in this case is due to local overheating of the film above the critical temperature T_c near such defects. In this case the formation of a finite-size normal-phase region (stable temperature domain) localized at a defect can precede microwave breakdown of an HTSC film.[19]

In summary, there exist two different scenarios leading to the development of microwave breakdown of a film and each scenario has its own characteristic spatiotemporal pattern of destruction of superconductivity. In the present paper we examine the interrelation of these scenarios and obtain breakdown criteria for both cases.

CRITICAL THERMAL DISTURBANCES

Let us consider an HTSC film with thickness D_f on a dielectric substrate of thickness $D_s \gg D_f$. The temperature of the back of the substrate is stabilized at T_0 (Fig. 1). Incident microwave radiation with power density P is partially absorbed in the film and the absorption in the substrate is negligibly small.[11,12] The power density P of the incident radiation is related with the surface microwave field B as
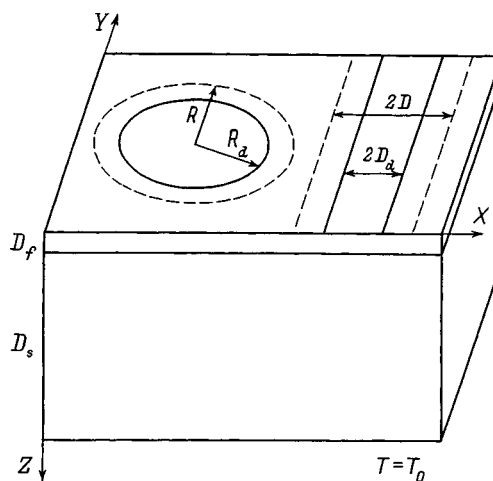

FIG. 1. Geometry of the problem and diagram of normal-phase nucleation on defects. Regions of the HTSC film — the regions bounded by solid lines are defects; the regions bounded by dashed lines are temperature domains.

$$P = \frac{B^2}{2\varepsilon_0^{1/2}\mu_0^{3/2}}. \tag{1}$$

Let the HTSC film be subjected initially at time $t=0$ to an external thermal disturbance of the energy

$$E = \int_{-\infty}^{\infty} dX \int_{-\infty}^{\infty} \int_0^{D_f} C_f T(X,Y,Z,0) dZ, \tag{2}$$

where X, Y, and Z are coordinates (Fig. 1), C_f is the specific heat of the film, and $T(X,Y,Z,0)$ is the initial temperature distribution in the film.

The destruction of superconductivity by a thermal disturbance is of a threshold character.[20] The critical energy corresponds to the minimum energy $E=E_c$ above which the normal phase arising after the action of a heat pulse propagates over the entire film. The critical energy is a complicated function of the initial temperature distribution (spatial extent of the disturbance[20]), so that two limiting cases can be conveniently distinguished from the entire spectrum of possible disturbances: local disturbances and spatially extended (nonlocal) disturbances.

Local thermal disturbances. We shall consider first the destruction of superconductivity by a local pulsed disturbance (heat pulse). The duration and extent of the heat pulse are small compared with the characteristic spatial and temporal scales of the problem (see below), so that the initial temperature distribution can be described by means of a delta function.[20] For typical ratios of the parameters of the film–substrate system $k_f D_s \gg k_s D_f$ (k_f and k_s are the thermal conductivity of the film and substrate, respectively) the temperature of the film is uniform over the film thickness,[13-15] and the heating of the film by the microwave radiation after the action of the pulse is described by the equation

$$C_f \frac{\partial T}{\partial t} = \frac{\partial}{\partial X}\left(k_f \frac{\partial T}{\partial X}\right) + \frac{\partial}{\partial Y}\left(k_f \frac{\partial T}{\partial Y}\right) + \frac{k_s}{D_f}\left(\frac{\partial T}{\partial Z}\right)\Big|_{Z=D_f}$$
$$+ \frac{\kappa(T)P}{D_f} + \frac{E}{D_f}\delta(X)\delta(Y)\delta(t), \quad 0<Z<D_f. \tag{3}$$

Here $\kappa(T)$ is the radiation absorption coefficient of the film, $\delta(X)$ is a delta function, the third term on the right-hand side of Eq. (3) describes the removal of heat from the film into the substrate, and the last term corresponds to the specific power of the heat pulse.[20]

The temperature dependence of the radiation absorption coefficient in the microwave range can be approximated by a step function[13-17,19]

$$\kappa(T) = \kappa_n \eta(T-T_c), \tag{4}$$

where $\kappa_n = (\sigma D_f Z_0)^{-1}$ and σ are, respectively, the normal-state absorption coefficient and conductivity of the film, $Z_0 = (\mu_0/\varepsilon_0)^{1/2} = 377 \ \Omega$ is the characteristic impedance of free space, $\eta(x)$ is the Heaviside step function ($\eta=0$ for $x<0$ and $\eta=1$ for $x\geq 0$).

The temperature distribution in the substrate and the boundary condition on the backside with respect to the film are described by the equations

$$C_s \frac{\partial T}{\partial t} = \frac{\partial}{\partial X}\left(k_s \frac{\partial T}{\partial X}\right) + \frac{\partial}{\partial Y}\left(k_s \frac{\partial T}{\partial Y}\right) + \frac{\partial}{\partial Z}\left(k_s \frac{\partial T}{\partial Z}\right),$$
$$D_f < Z < D_f + D_s, \tag{5}$$

$$T=T_0, \quad Z=D_f+D_s, \tag{6}$$

where C_s is the specific heat of the substrate.

Assuming for simplicity that the temperature dependences of C_s and k_s can be neglected and taking account of the fact that the film thickness is small compared with the substrate thickness ($D_f \ll D_s$), we represent Eqs. (5) and (6) in the form

$$\frac{\partial\theta}{\partial\tau} = \frac{\partial^2\theta}{\partial\rho^2} + \frac{1}{\rho}\frac{\partial\theta}{\partial\rho} + \frac{\partial^2\theta}{\partial z^2}, \quad 0<z<1, \tag{7}$$

$$\theta=0, \quad z=1. \tag{8}$$

Here $\rho = (X^2+Y^2)^{1/2}/D_s$ and $z=Z/D_s$ are dimensionless radial and transverse coordinates, $\tau = t k_s/C_s D_s^2$ is the dimensionless time, and $\theta=(T-T_0)/(T_c-T_0)$ is the dimensionless temperature. For an HTSC film on a dielectric substrate, the typical ratios of the parameters are $C_f D_f \ll C_s D_s$ and $k_f D_f \ll k_s D_s$.[14-17] This makes it possible to neglect the first three terms in Eq. (3) and, taking account of Eqs. (1) and (4), to represent Eq. (3) as a boundary condition for Eq. (7)

$$\frac{\partial\theta}{\partial z} + 2(B/B_b)^2\eta(\theta-1) = 0, \quad z=0. \tag{9}$$

Here B_b is the threshold breakdown field of a uniform film, determined by the expression

$$B_b = 2\mu_0[k_s\sigma D_f(T_c-T_0)/D_s]^{1/2}. \tag{10}$$

According to Eq. (1), B_b corresponds to the threshold breakdown power density of a uniform film $P_b = Z_0 B_b^2/2\mu_0^2 = 2k_s\sigma D_f Z_0(T_c-T_0)/D_s$.[13] For a typical Y–Ba–Cu–O film ($D_f \simeq 5\times10^7$ m, $\sigma \simeq 2\times10^6 \ \Omega^{-1}\cdot$ m^{-1}, $T_c \simeq 90$ K) on an LaAlO$_3$ substrate ($k_s \simeq 20$ W$\cdot$m$^{-1}\cdot$K^{-1}) of thickness $D_s \simeq 5\times10^{-4}$ m, at $T_0 \simeq 77$ K we obtain $B_b \approx 2$ mT and $P_b \approx 4\times10^4$ W$\cdot$cm^{-2}. The initial condition for Eq. (7) is

$$\theta(\rho,z,0) = \frac{e}{2\pi\rho}\delta(\rho)\delta(z). \tag{11}$$

Here $e=E/E_h$ is the dimensionless energy of the heat pulse and $E_h=C_s D_s^3(T_c-T_0)$ is a characteristic "thermal" energy. For a typical Y–Ba–Cu–O film on an LaAlO$_3$ substrate ($C_s \simeq 10^5$ J$\cdot$cm$^{-3}\cdot$K^{-1}) of thickness $D_s \simeq 5\times10^{-4}$ m, we obtain $E_h \approx 1.5\times10^{-4}$ J. It follows from Eqs. (7)–(9) and (11) that the dimensionless critical energy $e_c = E_c/E_h$ depends only on the dimensionless surface microwave field B/B_b. Thus dimensional analysis shows that the characteristic spatial and temporal scales of the problem are D_s and $C_s D_s^2/k_s$, respectively, and it makes it possible to obtain the dependence of the critical energy E_c on all parameters except the surface microwavefield.

The problem (7)–(9) and (11) cannot be solved analytically. But e_c can be estimated on the basis of the following considerations. Let a region of radius R_0 in the film be transferred into the normal state immediately after the action of
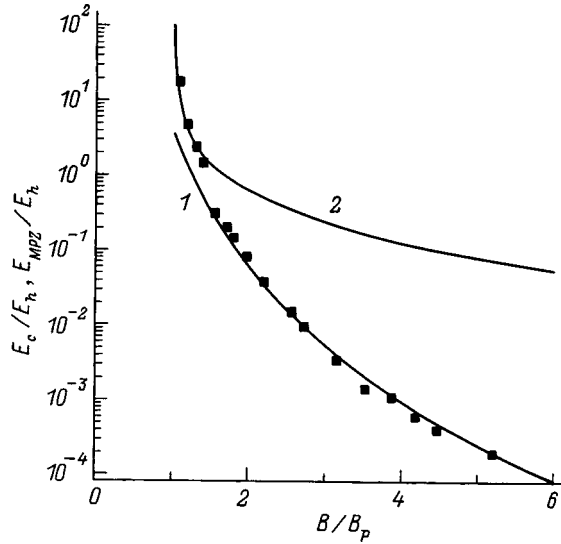
FIG. 2. $E_c(B)$ (*1* — Eq. (12), ■ — numerical calculation) and $E_{MPZ}(B)$ (*2*).

the critical pulse. Correspondingly, a hemisphere of the same radius is heated in the substrate. Then $e_c \sim (2\pi/3)\Delta\theta r_0^3$, where $\Delta\theta$ is the characteristic temperature increase in the heated region and $r_0 = R_0/D_s$. The quantity r_0 can be estimated in order of magnitude by equating the heat release in the normal region of the film $\sim 2(B/B_b)^2\pi r_0^2$ to the removal of heat into the ''cold'' part of the substrate $\sim (\Delta\theta/r_0)2\pi r_0^2$. From this equation, taking for the film $\Delta\theta \sim 1$, we obtain $r_0 \sim (B/B_b)^{-2}$. Therefore the energy of a critical disturbance destroying superconductivity in the film is $e_c \approx (2\pi/3) \times (B/B_b)^{-6}$. This relation holds if the radius of the heated region is small compared with the film thickness ($r_0 \ll 1$). Therefore the estimate obtained describes the asymptotic behavior of e_c for $B \gg B_b$.

To check the accuracy and range of applicability of the relation $e_c \propto (B/B_b)^{-6}$ obtained above, the destruction of superconductivity was simulated numerically. A Gaussian distribution was chosen for the initial temperature profile in the film–substrate system. It was established that the numerically obtained value of the critical energy e_c does not depend on the half-width of the initial temperature profile, if the locality condition holds (the half-width of the initial temperature profile is small compared with the film thickness). The function $E_c(B)$ obtained by solving Eq. (7) numerically with the boundary conditions (8) and (9) is shown in Fig. 2. It follows from Fig. 2 that the asymptotic estimate obtained above describes accurately the functional dependence $E_c(B)$ for $B \gg B_p$ and can be represented as

$$E_c/E_h = 4.0(B/B_b)^{-6}. \tag{12}$$

Extended thermal disturbances. The superconducting state in a film can also be destroyed by extended (nonlocal) disturbances, whose range of action is comparable in size to the thickness of the film. The critical energy of such disturbances can be estimated as the formation energy of a ''minimum propagating zone,''[20,21] playing the role of a critical nucleus of the normal phase. The temperature distribution in the ''minimum propagating zone'' $\theta = \theta_{MPZ}(\rho,z)$ is a sta-

tionary ($\partial\theta/\partial\tau = 0$) unstable solution of Eq. (7), satisfying the boundary conditions (8) and (9) and $\partial\theta/\partial\rho = 0$ at $\rho = 0$ and $\rho = \infty$.[22,23] Such a solution can be found by separation of variables,[15] which gives for the formation energy of the ''minimum propagating zone''

$$e_{MPZ} = 2\pi \int_0^\infty \rho\, d\rho \int_0^1 \theta_{MPZ}(\rho,z)dz \tag{13}$$

the expression

$$4q \sum_{k=0}^\infty \lambda_k^{-1} I_1(\lambda_k q) K_0(\lambda_k q) = (B/B_b)^{-2}. \tag{14}$$

Here $\lambda_k = \pi(2k+1)/2$, $q = (B_b/B)(e_{MPZ}/\pi)^{1/2}$, and $I_1(x)$ and $K_0(x)$ are modified Bessel functions. The function $E_{MPZ}(B)$ calculated from Eq. (14) is shown in Fig. 2. Using the Euler–Maclaurin summation formula we obtain from Eq. (14) $E_{MPZ}/E_h \simeq (\pi^3/16)(B/B_b)^{-2}$ for $B \gg B_b$.

In summary, in a wide range of values of the surface microwave field $B \gg B_b$ the critical energy of heat pulses ($E_c \propto B^{-6}$) is less than the formation energy of the ''minimum propagating zone'' ($E_{MPZ} \propto B^{-2}$). This means that heat pulses with a lower critical energy are most ''dangerous'' from the standpoint of destroying superconductivity.

MICROWAVE BREAKDOWN AT DEFECTS

Let us consider an HTSC film with a normal (nonsuperconducting) defect, whose absorption coefficient $\kappa = \kappa_d$ is temperature-independent. When the film in the superconducting state is exposed to microwave radiation, absorption first occurs only on a defect, as a result of which the temperature of the defect increases. The region of the film and substrate adjoining the defect is heated as a result of heat diffusion. If the surface microwave field exceeds the threshold value, then the region of the film adjoining the defect is transferred into the normal state and becomes a source of additional absorption of microwave radiation. Subsequently, the normal-phase region arising either propagates over the entire film or a finite-size normal-phase region (a stable temperature domain) forms near the defect. The temperature distribution in the domain depends on the geometric shape of the defect and can be found analytically for circular and linear defects.

Circular defect. Let us consider first an HTSC film with a circular defect of radius R_d (Fig. 1). In this case the boundary condition for Eq. (7) at $z = 0$ is

$$\frac{\partial\theta}{\partial z} + 2(B/B_b)^2\{\gamma\eta(r_d-\rho) + \eta(\theta-1)\eta(\rho-r_d)\} = 0, \; z = 0, \tag{15}$$

where $r_d = R_d/D_s$ is the dimensionless radius of the defect and $\gamma = \kappa_d/\kappa_n$ is the ratio of the absorption coefficients of the defect and the film in the normal state, characterizing the ''contrastness'' of the defect.

The temperature distribution $\theta(\rho,z)$ in the domain, satisfying the stationary Eq. (7) ($\partial\theta/\partial\tau = 0$) with the boundary conditions (8) and (15) and $\partial\theta/\partial\rho = 0$ at $\rho = 0$ and $\rho = \infty$, can be found by separation of variables. For domain radius R,
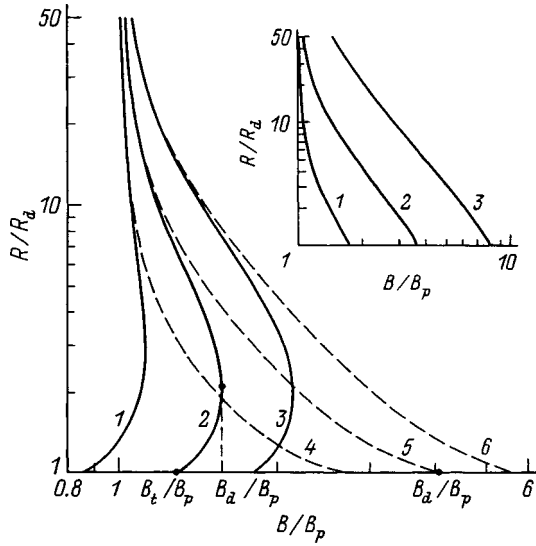
FIG. 3. Circular defect. The domain radius versus the surface microwave field: $\gamma=5$ (solid curves), 0.5 (dashed curves); R_d/D_s: *1,4 — 0.25; 2,5 — 0.1; 3,6 — 0.05.* Inset: $R(B)$ for $\gamma=1.2$. R_d/D_s: *1 — 0.25, 2 — 0.05, 3 — 0.01.*

determined from the condition $\theta(r,0)=1$ ($r=R/D_s$ is the dimensionless radius of the domain), we obtain the relation

$$(R/D_s)\sum_{k=0}^{\infty}\lambda_k^{-1}I_1(\lambda_k R/D_s)K_0(\lambda_k R/D_s)$$

$$+(R_d/D_s)(\gamma-1)\sum_{k=0}^{\infty}\lambda_k^{-1}I_1(\lambda_k R_d/D_s)$$

$$\times K_0(\lambda_k R/D_s)=(B_p/2B)^2. \qquad (16)$$

We note that a dependence similar to Eq. (16) was obtained in Ref. 19. The dependence $R(B)$ is presented in Fig. 3 for various values of R_d for the cases $\gamma>1$ (solid lines) and $\gamma<1$ (dashed lines). It is evident from Fig. 3 that the character of the destruction of superconductivity depends strongly on the ''contrastiness'' γ of the defect.

We shall examine in greater detail the spatiotemporal picture of breakdown for the case $\gamma>1$. As the surface microwave field B increases, the temperature of the defect increases, but the region of the film next to the defect remains superconducting. Above a certain field $B=B_t$, a normal phase ring forms around the defect. The threshold field B_t for the formation of a domain is determined from the condition $R(B_t)=R_d$. Using this condition, from Eq. (16) we obtain for B_t

$$B_t/B_b$$

$$=\left[4\gamma(R_d/D_s)\sum_{k=0}^{\infty}\lambda_k^{-1}I_1(\lambda_k R_d/D_s)K_0(\lambda_k R_d/D_s)\right]^{-1/2}.$$

$$(17)$$

Simple analytic expressions for B_t can be obtained from Eq. (17) in the limiting cases of defects with small radii $R_d \ll D_s$ and large radii $R_d \gtrsim D_s$. According to the Euler–

Maclaurin formula, for $R_d \ll D_s$ the sum in Eq. (17) can be replaced by a corresponding integral. In this approximation we have for B_t

$$B_t/B_b=\frac{1}{2}\left[\frac{\pi}{\gamma R_d/D_s}\right]^{1/2}, \quad R_d \ll D_s. \qquad (18)$$

For $R_d \gtrsim D_s$ the first term in the sum in Eq. (17) is sufficient to obtain an asymptotic expression for B_t. Then we obtain for B_t

$$B_t/B_b=\gamma^{-1/2}\left[1-\frac{7\xi(3)}{\pi^3 R_d/D_s}\right]^{-1/2}, \quad R_d \gtrsim D_s. \qquad (19)$$

Here $\xi(3)$ is the Riemann zeta function. It follows from Eq. (19) that for $\gamma>1$ the threshold field B_t for domain formation can be both greater and less than the threshold breakdown field B_b of a uniform film.

For $B>B_t$ the radius R of the temperature domain arising on a defect increases with the surface microwave field (curves *1–3* in Fig. 3). Linear analysis of the stability of the solution $\theta(\rho,z)$ with respect to small disturbances shows that the branch with $dR/dB>0$ corresponds to a stable domain.[24] A stable domain vanishes at a certain value of the surface microwave field $B=B_d$ (Fig. 3). The condition for vanishing of a stable domain corresponds to microwave breakdown of the entire film.[19] Indeed, it is evident from Fig. 3 that for any defect radius the breakdown field B_d is always greater than the threshold breakdown field B_b of a uniform film. Therefore for $B>B_d$ the normal-phase region, arising on a defect propagates over the entire sample. Approximate analytic expressions for the breakdown field B_d can be obtained in the limiting cases of defects with a small radius $R_d \ll D_s$ and defects with a large radius $R_d \gtrsim D_s$

$$\frac{B_d}{B_b}=\frac{\pi^{1/4}}{2(\gamma-1)^{1/4}(R_d/D_s^{1/2})}, \quad R_d \ll D_s, \qquad (20)$$

$$\frac{B_d}{B_b}=\left[1-\frac{\alpha}{\ln(\beta/\alpha)}+\frac{\alpha}{\ln^2(\beta/\alpha)}\right]^{-1/2}, \quad R_d \gtrsim D_s. \qquad (21)$$

Here $\alpha=7\xi(3)/2\pi^2$ and

$$\beta=2(\gamma-1)(R_d/D_s)^{1/2}\exp(\pi R_d/2D_s).$$

In the case $\gamma<1$ a stable temperature domain does not arise in the film for any defect radius (curves *4–6* in Fig. 3) and microwave breakdown occurs for $B_d=B_t$. It should be noted that even in the case $\gamma>1$, for sufficiently small defect radii R_d the solution corresponding to a stable domain vanishes (inset in Fig. 3), and similarly to the case $\gamma<1$ breakdown occurs at $B_d=B_t$, bypassing the intermediate stage of formation of a stable temperature domain in the film. Figure 4 compares the results of the numerical calculation of B_d using the transcendental equation (16) with the analytic formulas (18)–(21). It follows from Fig. 4 that the approximate expressions (18)–(21) agree well with the numerical results.
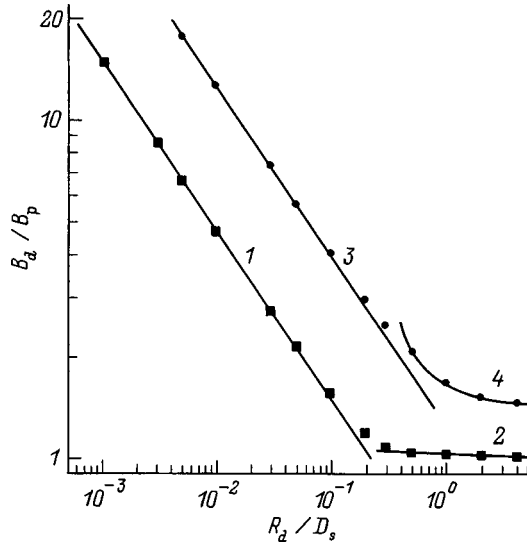
FIG. 4. Circular defect: breakdown field versus the defect radius. $\gamma=5$ (1 — Eq. (20), 2 — Eq. (21), ■— numerical calculation); $\gamma=0.5$ (3 — Eq. (18), 4 — Eq. (19), ● — numerical calculation).

FIG. 5. Linear defect. The domain half-width versus the surface magnetic field: $\gamma=2$ (solid curves), 0.8 (dashed curves); D_d/D_s: 1,4 — 0.5; 2,5 — 0.1; 3,6 — 0.02. Inset: $D(B)$ for $\gamma=5$. D_d/D_s: 1 — 0.05, 2 — 0.02.

Linear defect. Let us now consider an HTSC film with a linear defect of half-width D_d (Fig. 1). In this case the stationary temperature distribution $\theta(x,z)$ in the substrate is described by the two-dimensional equation

$$\frac{\partial^2 \theta}{\partial x^2} + \frac{\partial^2 \theta}{\partial z^2} = 0, \quad 0 < z < 1, \tag{22}$$

where $x = X/D_s$.

By virtue of the symmetry of the problem we shall consider below only the region $x > 0$. In this region the boundary condition to Eq. (22) at $z = 0$ can be represented as

$$\frac{\partial \theta}{\partial z} + 2(B/B_b)^2\{\gamma\eta(d_d-x) + \eta(\theta-1)\eta(x-d_d)\} = 0, \quad z=0, \tag{23}$$

where $d_d = D/D_s$ is the dimensionless half-width of the defect. The equation (22) with the boundary conditions (8) (23) and $\partial\theta/\partial x = 0$ at $x = 0$ and $x = \infty$ can be solved by separation of variables, which gives for the domain half-width D, determined from the condition $\theta(d,0) = 1$ ($d = D/D_s$ is the dimensionless domain half-width) the expression

$$\sum_{k=0}^{\infty} \lambda_k^{-2} \exp(-2\lambda_k D/D_s) - 2(\gamma-1)$$

$$\times \sum_{k=0}^{\infty} \lambda_k^{-2} \sinh(\lambda_k D_d/D_s) \exp(-\lambda_k D/D_s)$$

$$= (B^2 - B_b^2)/2B^2. \tag{24}$$

Figure 5 shows the domain half-widths D calculated using expression (24) as a function of the surface microwave field B for different values of the defect half-width D_d for $\gamma > 1$ (solid curves) and $\gamma < 1$ (dashed curves). It is evident from Fig. 5 that for $\gamma < 1$ the function $D(B)$ (curves 4–6 in Fig. 5) is qualitatively similar to the function $R(B)$ for a circular defect. In this case microwave breakdown occurs in
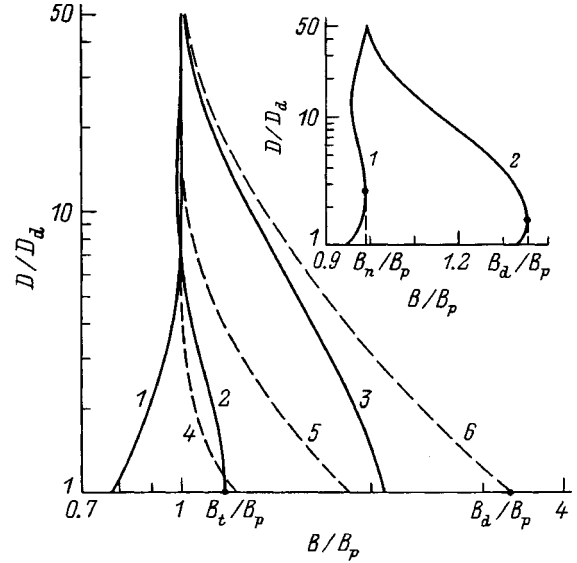
a manner by passing the intermediate stage of formation of a stable temperature domain in the film, and the breakdown field B_d equals the threshold field B_t for the formation of a normal phase in the film. The quantity B_t is determined from the condition $R(B_t) = D_d$, which together with Eq. (24) gives

$$\frac{B_t}{B_b} = \gamma^{-1/2} \left[1 - 2\sum_{k=0}^{\infty} \lambda_k^{-2} \exp(-2\lambda_k D_d/D_s) \right]^{-1/2}. \tag{25}$$

Approximate expressions for B_t can be obtained from Eq. (25) in the limiting cases of a small defect half-width $D_d \ll D_s$ and large defect half-width $D_d \gtrsim D_s$

$$\frac{B_t}{B_b} = \frac{1}{2} \left[\frac{\pi}{\gamma(D_d/D_s)|\ln(D_d/D_s)|} \right]^{1/2}, \quad D_d \ll D_s, \tag{26}$$

$$\frac{B_t}{B_b} = \gamma^{-1/2}[1 - (8/\pi^2) \exp(\pi D_d/D_s)]^{-1/2},$$

$$D_d \gtrsim D_s. \tag{27}$$

It follows from Eq. (27) that for $\gamma < 1$ B_t is always greater than the threshold breakdown field B_b of a uniform film, while for $\gamma > 1$ the threshold field B_t for the formation of a domain can be both greater and less than B_b.

For $\gamma > 1$ and $B_t < B_b$ a localized temperature domain forms near a defect with a sufficiently large half-width D_d. Its half-width increases monotonically with B and becomes infinite at $B = B_b$ (curve 1 in Fig. 5). In this case the breakdown field B_d equals the threshold breakdown field B_b of a uniform film: $B_d = B_b$. For smaller values of D_d a stable domain (branch with $dD/dB > 0$) vanishes at some field $B_n < B_b$ (curve 1 in the inset in Fig. 5). For $B > B_n$ a second, large, stable domain is formed. In this case breakdown also occurs for $B_d = B_b$. For even smaller values of D_d a stable domain vanishes at a breakdown field B_d greater than B_b (curve 2 in the inset in Fig. 5). In this case a second stable
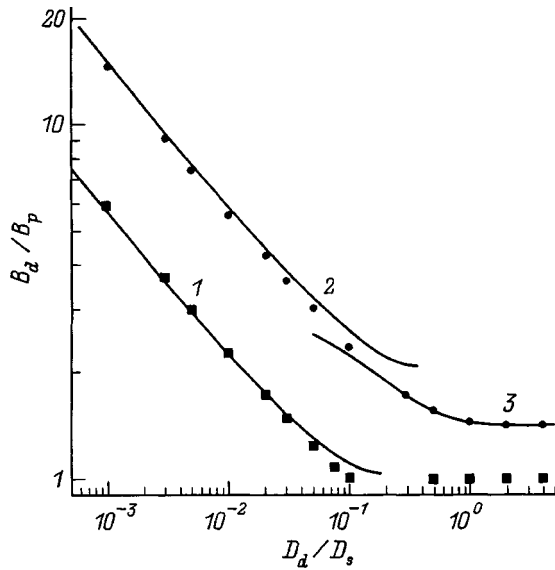
FIG. 6. Linear defect. The breakdown field versus the defect half-width. $\gamma=5$ (1 — Eq. (28), ■ —numerical calculation); 0.5 (2 — Eq. (26), 3 — Eq. (27), ● — numerical calculation).

domain with a larger half-width does not form, and for $B > B_d$ the film passes into the normal state as a result of the propagation of a normal phase arising at a defect over the entire region. An approximate analytic expression for B_d can be obtained in the limiting case of a small defect half-width $D_d \ll D_s$

$$\frac{B_d}{B_b} = \frac{1}{2}\left[\frac{\pi}{2(\gamma-1)(D_d/D_s)\left|\ln\{(\gamma-1)(D_d/D_s)\}\right|}\right]^{1/2},$$

$$D_d \ll D_s. \tag{28}$$

The computational results, obtained using the asymptotic formulas (26)–(28), for the breakdown field B_d are compared in Fig. 6 with the numerical calculations performed using the transcendental equation (24). It is evident that for $\gamma > 1$ Eq. (28) satisfactorily describes the function $B_d(D_d)$ right up to $D_d \approx D_s$, while for $D_d > D_s$ the breakdown field B_d equals the threshold breakdown field B_b of a uniform film. It should be noted that similarly to the case of a circular defect for $\gamma > 1$ and sufficiently small defect half-width D_d, a stable domain is not formed in the film (curve 3 in Fig. 5). In this case, just as for $\gamma < 1$, microwave breakdown occurs for $B_d = B_t$.

DISCUSSION

In summary, the thermal mechanism of microwave breakdown of an HTSC film is due to strong dissipation in the normal region, which under certain conditions can result in the propagation of a normal phase over the entire film.[13–16] Such propagation is possible under two conditions. First, the surface microwave field B must be greater than the threshold breakdown field of a uniform film $B_b = 2\mu_0[k_s\sigma D_f(T_c - T_0)/D_s]^{1/2}$. Second, a sufficiently large normal-phase region must arise in the film. Two scenarios for the appearance of a normal-phase region sufficient for

microwave breakdown of a film are possible. Each scenario has its own spatiotemporal pattern of destruction of superconductivity in the film.

The first scenario of the formation of a normal-phase region in the film is due to the action of external thermal disturbances, which are always present in real systems. For $B > B_b$ thermal disturbances with energy greater than the critical value E_c lead to the appearance of a normal-phase region sufficient for subsequent propagation of the phase over the entire film. Such propagation occurs by unbounded expansion of the normal-phase region, whose boundary is a temperature autowave of S–N switching.[13–16] The critical energy depends strongly on the spatial extent of the disturbances and decreases with increasing surface microwave field. For qualitative analysis it is convenient to separate from the spectrum of possible disturbances two limiting cases: local disturbances and spatially extended disturbances. The critical energy of spatially extended disturbances can be estimated as the formation energy of the ''minimum propagating zone'' E_{MPZ},[21] which plays the role of a critical nucleus of the normal phase. In a wide range of values of the surface microwave field $B \gg B_b$ the critical energy of local disturbances ($E_c \propto B^{-6}$) is much lower than the energy of spatially extended disturbances ($E_{MPZ} \propto B^{-2}$). Therefore local disturbances are most ''dangerous'' from the standpoint of the destruction of superconductivity, since the transition into the normal state is initiated mainly by disturbances with the minimum critical energy.[20]

The second scenario of microwave breakdown is due to the presence of nonsuperconducting defects in an HTSC film. Local overheating near such defects can destroy superconductivity. In this case a stage of formation of a finite-size normal-phase region (a stable temperature domain) at a defect can precede microwave breakdown. We note that local destruction of superconductivity in HTSC films has been observed in a number of experiments.[18,25,26] The spatiotemporal pattern of destruction of superconductivity and the breakdown field B_d depend strongly on the ratio γ of the absorption coefficients of the defect and the film in the normal state (the ''contrastiness'' of a defect). For $\gamma > 1$, as a rule, a stage of formation of a temperature domain localized on a defect precedes breakdown, while for $\gamma < 1$ breakdown occurs without the stage of formation of a stable domain. The breakdown field B_d also depends on the form of the defect. For a circular defect the breakdown field B_d is always greater than the threshold breakdown field B_b of a uniform film, and for a linear defect B_d can be both greater than and equal to B_b. The breakdown field B_d increases with decreasing defect size: $B_d \propto (R_d/D_s)^{-1/2}$ for a circular defect and $B_d \propto [D_d/D_s|\ln(D_d/D_s)|]^{-1/2}$ for a linear defect. It follows from these relations, specifically, that for comparable scales of defect sizes the breakdown field on a linear defect can be several times lower than on a circular defect. Therefore linear defects are most dangerous from the standpoint of the destruction of superconductivity.

Real HTSC films contain a variety of defects of various sizes. As follows from the relations obtained above, microwave breakdown will occur on the largest defect in a film. Comparing the analytic expressions obtained for B_d with the

experimental data[1–8] shows that the idea of overheating of the film by microwave radiation near defects of size R_d, $D_d \simeq 1-10$ μm makes it possible to describe satisfactorily the observed values of the breakdown fields ($B_d \sim 10$ mT). However, as the quality of HTSC films improves (the defect sizes decreases), the breakdown field B_d probably will not increase appreciably. This is due to the fact that for high surface microwave fields external disturbances which are always present in real systems have a very low critical energy. For example, it follows from Eq. (12) that for typical values of the film and substrate parameters presented above the critical energy for $B > 10$ mT is very low $E_c < 10^{-8}$ J. It is obvious that near small defects breakdown will occur much earlier than overheating because of the destruction of superconductivity of uniform regions in the film by external disturbances. Therefore, depending on the ratio of the energy of the disturbances acting in the system and the size of the defects in the film, one of the two scenarios of the development of microwave breakdown which were studied above will predominate.

[1] R. W. Ralston, Supercond. Sci. Technol. **4**, 386 (1991).

[2] W. L. Holstein, L. A. Parisi, Z.-Y. Shen *et al.*, J. Supercond. **6**, 191 (1993).

[3] W. Diete, B. Aschermann, H. Chaloupka *et al.*, Appl. Supercond. **2**, 1107 (1995).

[4] S. Hensen, M. Lenkens, M. Getta *et al.*, Appl. Supercond. **2**, 1127 (1995).

[5] M. Manzel, S. Huber, H. Bruchlos *et al.*, Appl. Supercond. **2**, 1155 (1995).

[6] N. Klein, N. Tellmann, U. Dähne *et al.*, IEEE Trans. Appl. Supercond. **AS-5**, 2663 (1995).

[7] W. Diete, M. Getta, M. Hein *et al.*, IEEE Trans. Appl. Supercond. **AS-7**, 1236 (1997).

[8] T. Kaiser, C. Bauer, and W. Diete, Appl. Supercond. **1**, 45 (1997).

[9] J. Halbritter, IEEE Appl. Supercond., 662 (1972).

[10] H. Padamsee, IEEE Trans. Magn. **MAG-19**, 1322 (1983).

[11] A. A. Zharov, A. L. Korotkov, and A. N. Reznik, Sverkhprovodimost' (KIAE) **5**(3), 419 (1992).

[12] A. A. Zharov, A. L. Korotkov, and A. N. Reznik, Supercond. Sci. Technol. **5**, 104 (1992).

[13] A. A. Pukhov, Pis'ma Zh. Tekh. Fiz. **22**(12), 55 (1996) [Tech. Phys. Lett. **22**(6), 504 (1996)].

[14] N. A. Buznikov and A. A. Pukhov, Pis'ma Zh. Tekh. Fiz. **22**(12), 45 (1996) [Tech. Phys. Lett. **22**(6), 499 (1996)].

[15] N. A. Buznikov and A. A. Pukhov, Supercond. Sci. Technol. **10**, 318 (1997).

[16] A. A. Zharov, I. M. Nefedov, and A. N. Reznik, Zh. Tekh. Fiz. **67**(10), 81 (1997) [Tech. Phys. **42**, 1184 (1997)].

[17] N. A. Buznikov and A. A. Pukhov, Pis'ma Zh. Tekh. Fiz. **22**(20), 51 (1996) [Tech. Phys. Lett. **22**(10), 843 (1996)].

[18] M. Hein, W. Diete, and M. Getta *et al.*, IEEE Trans. Appl. Supercond. **AS-7**, 1264 (1997).

[19] A. A. Zharov and A. N. Reznik, Zh. Tekh. Fiz. **68**(1), 131 (1998) [Tech. Phys. **43**, 117 (1998)].

[20] A. Vl. Gurevich, R. G. Mints, and A. L. Rakhmanov, *The Physics of Composite Superconductors* (Nauka, Moscow, 1987, 240 pp.).

[21] A. A. Pukhov, Sverkhprovodimost' (KIAE) **6**(2), 235 (1993).

[22] A. P. Martinelli and S. L. Wipf, IEEE Appl. Supercond., 325 (1972).

[23] M. N. Wilson and Y. Iwasa, Cryogenics **18**, 17 (1978).

[24] A. Vl. Gurevich and R. G. Mints, Usp. Fiz. Nauk **142**, 61 (1984) [Sov. Phys. Usp. **27**, 19 (1984)].

[25] G. Hampel, P. Kolodner, P. L. Gammel *et al.*, Appl. Phys. Lett. **69**, 571 (1996).

[26] J. Wosik, L.-M. Xie, J. H. Miller Jr. *et al.*, IEEE Trans. Appl. Supercond. **AS-7**, 1470 (1997).

Translated by M. E. Alferieff

Measure of the closeness of crystal lattices

V. P. Vereshchagin and M. P. Kashchenko

Urals State Forest Technology Academy, 620032 Ekaterinburg, Russia

A numerical measure of the difference between crystal lattices is determined. The fruitfulness of the definition is demonstrated for a specific example concerning the prominence of an orientational correspondence between the body-centered crystal lattice [bcc(bct)] of α-martensite and the face-centered crystal lattice (fcc) of γ-austenite in cases where the latter is perfect and where it differs from a perfect lattice near a screw dislocation. © *1999 American Institute of Physics.* [S1063-7842(99)01005-3]

1. In the theory of martensitic transformations (MTs) ideas associating the possibility of a spontaneous structural rearrangement of austenite to definite locations where the austenite has been deformed beforehand are used in the discussion of the nucleation stage (see, for example, Refs. 1–3). It is believed that these deformations bring the austenite and martensite lattices into coincidence, thereby providing favorable conditions for fluctuation activation above an energy barrier. The results of Refs. 4 and 5, for example, provide evidence of this. In Ref. 4 it was established that there exists a temperature M exceeding the MT onset temperature M_s by 35–40 K. At temperatures in the interval (M_S, M) the MT is produced by external stresses $\sigma_m < \sigma_\tau$, where σ_τ is the yield point. Strains $\varepsilon < 10^{-3}$ correspond to such values of σ_m. Therefore it can be concluded that elastic strains produced by an external load decrease the martensite nucleation threshold. An orientational effect is described in Ref. 5 — the appearance of predominantly large-angle orientations of athermal martensite plates under elastic tension, indicating that uniaxial elastic stresses act selectively, lowering the barrier for nucleation of martensite crystals of some orientations and raising it for crystals with all other orientations. Analysis of the effect[6] shows that the barrier is decreased if the stresses lead to compression and tension, promoting the Bain variants of deformation, which occur in a martensite plate, and increased otherwise. On the other hand, the effect of a dislocation, for example, is in no way inferior in this respect to an external load and therefore could decrease the barrier for martensite nucleation at some location or locations. Moreover, Refs. 7–10 attest (with respect to the $\gamma \to \alpha$ MT in iron alloys at least) to the fact that dislocations are involved in the martensite nucleation and that a spontaneous MT is impossible in perfect austenite.

At the same time it is obvious that at the level of elastic strains a lattice restructuring, such as, let us say, fcc $\to$ bcc(bct), by the Bain scheme is impossible. Elastic compression and tension can only bring the fcc and bcc(bct) lattices into closer coincidence, giving rise to Bain deformation. This leads to difficulties, making it necessary to construct a picture of elastic deformations that lower the barrier for martensite nucleation, since the concept of the closeness of lattices is itself, even though implicitly postulated in a number of approaches, nonquantitative. A quantitative characterization of the closeness of lattices requires a measure that expresses the difference between the lattices in terms of a number. In principle such a measure is also necessary to study many questions concerning the conjunction of regions coexisting in a heterogeneous solid and having different crystal lattices. The principal objective of the present paper is to determine a measure of the closeness of lattices. The fruitfulness of the proposed definition is illustrated here for a specific example which is of interest for assessment of the effect of an elastic field of a screw dislocation on the prominence of an orientational correspondence between the bcc(bct) lattice of α-martensite and the fcc lattice of γ-austenite.

2. Let $\{\mathbf{X}\}$ and $\{\mathbf{X}'\}$ be lattices. The problem is to determine a measure characterizing quantitatively the difference between these lattices. Since we are talking about applications to martensitic transformations, it is possible and convenient to formulate the problem differently — in terms of mappings. Indeed, let the lattices $\{\mathbf{X}\}$ and $\{\mathbf{X}'\}$ be the result of deformations of a given lattice $\{\mathbf{r}\}$ which are described by the mappings

$$\mathbf{X}_i = \mathbf{L}\mathbf{r}_i, \quad \mathbf{X}'_r = \mathbf{T}\mathbf{r}_i; \quad i = 1,2,3, \tag{1}$$

where $\mathbf{X}_i$, $\mathbf{X}'_i$, and $\mathbf{r}_i$ are translational vectors of the lattices $\{\mathbf{X}\}$, $\{\mathbf{X}'\}$, and $\{\mathbf{r}\}$, respectively.

Then the closeness of the lattices $\{\mathbf{X}\}$ and $\{\mathbf{X}'\}$ will mean closeness of the tensors $\mathbf{L}$ and $\mathbf{T}$ and vice versa. The pair of tensors $\mathbf{L}$ and $\mathbf{T}$ can be associated to a corresponding number

$$m(\mathbf{L}, \mathbf{T}) = \{\text{Tr}[(\mathbf{L} - \mathbf{T})(\mathbf{L}^* - \mathbf{T}^*)]\}^{1/2}, \tag{2a}$$

which can be interpreted as the distance between the points representing the tensors $\mathbf{L}$ and $\mathbf{T}$ in an abstract nine-dimensional Euclidean space and satisfying the conditions $m(\mathbf{L}, \mathbf{L}) = 0$, $m(\mathbf{L}, \mathbf{T}) > 0$ $(\mathbf{L} \neq \mathbf{T})$, $m(\mathbf{L}, \mathbf{T}) = m(\mathbf{T}, \mathbf{L})$, $m(\mathbf{L}, \mathbf{T}) \leq m(\mathbf{L}, \mathbf{F}) + m(\mathbf{F}, \mathbf{T})$, which are the usual conditions for the distance between points. In Eq. (2a) and below an asterisk signifies transposition.

We note that Eq. (2a), which establishes a correspondence between numbers and tensor pairs, was obtained by generalizing the formula

$$m(A,B) = \left\{ \sum_{i=1}^{3} (a_i - b_i)^2 \right\}^{1/2},$$

which associates to each pair of points A and B in a three-dimensional Euclidean space a number $m(A,B)$, the distance between the points. The closeness or remoteness of the points can be judged according to the values of m. The position of the points A and B is given by the triplet of numbers $\{a_i\}$ and $\{b_i\}$ — their coordinates relative to a Cartesian rectangular coordinate system. The rank-2 tensors **L** and **T** in this coordinate system are represented by the sets $\{L_{in}\}$ and $\{T_{in}\}$ of matrix elements, where i, $n = 1$, 2, and 3, each set consisting of nine numbers. However, there is nothing wrong with considering the numbers as the coordinates of points (we denote them, just like the tensors, by **L** and **T**) in an abstract nine-dimensional Euclidean space and defining the distance between them as

$$m(\mathbf{L}, \mathbf{T}) = \left\{ \sum_{i,n=1}^{3} (L_{in} - T_{in})^2 \right\}^{1/2}$$

by analogy to the distance between points in a three-dimensional space. Hence it is easy to switch to the equation

$$m(\mathbf{L}, \mathbf{T}) = \left\{ \sum_{i=1}^{3} \sum_{n=1}^{3} (\mathbf{L} - \mathbf{T})_{in} (\mathbf{L}^* - \mathbf{T}^*)_{ni} \right\}^{1/2},$$

since $(L_{in} - T_{in})^2 = [(\mathbf{L} - \mathbf{T})_{in}]^2 = (\mathbf{L} - \mathbf{T})_{in} (\mathbf{L} - \mathbf{T})_{in}$ $= (\mathbf{L} - \mathbf{T})_{in} (\mathbf{L}^* - \mathbf{T}^*)_{ni}$, then to the equation

$$m(\mathbf{L}, \mathbf{T}) = \left\{ \sum_{i=1}^{3} [(\mathbf{L} - \mathbf{T})(\mathbf{L}^* - \mathbf{T}^*)]_{ii} \right\}^{1/2},$$

and finally to Eq. (2a), which does not depend explicitly on the choice of the coordinate system.

It is natural to take the number (2a), characterizing the difference between the tensors **L** and **T**, taking account of Eq. (1), as the characteristic of the difference between the lattices $\{\mathbf{X}\}$ and $\{\mathbf{X}'\}$ and to set

$$m(\{\mathbf{X}\}, \{\mathbf{X}'\}) = m(\mathbf{L}, \mathbf{T}). \tag{2b}$$

Then it will serve as a measure making it possible to express in a quantitative form the property of closeness of objects such as lattices and, together with concepts of equality (coincidence) and inequality (noncoincidence) of lattices, to establish an order relation for them. As a result it becomes possible to solve on a quantitative basis problems in which it is necessary to choose from a set of lattices the one that is closest to a given lattice. On the other hand the quantity $m(\{\mathbf{X}\}, \{\mathbf{r}\}) = m(\mathbf{L}, \mathbf{I})$, where **I** is the unit tensor, can be used as an order parameter for describing deformational phase transitions $\mathbf{L}\{\mathbf{r}\} \to \{\mathbf{X}\}$, which in contrast to the tensor **L** is a scalar. Indeed, m takes on different values: $m(\mathbf{I}, \mathbf{I}) = 0$ in the initial phase and $m(\mathbf{L}, \mathbf{I}) \neq 0$ in the final phase, characterizing the quantitative difference between the phases. Finally, we note with respect to the definition (2) that the abstract space of points representing tensors need not be Euclidean. A different metric can also be chosen (see, for example, Ref. 11), postulating, depending on the specific requirements of the problem, a metric tensor that is different from the unit tensor.

3. The formulation of the property of the closeness of lattices in terms of mappings leads to a formulation of the problem that is quite interesting from the standpoint of martensite problems. It arises in connection with the treatment of the mapping $\{\mathbf{r}\} \to \{\mathbf{X}\}$ as a sequence of mappings $\{\mathbf{r}\} \to \{\mathbf{X}'\} \to \{\mathbf{X}\}$ in accordance with the decomposition[12] of the tensor **L** into a product

$$\mathbf{L} = \mathbf{\Omega}\mathbf{E} \tag{3}$$

of an orthogonal tensor $\mathbf{\Omega}$ and a symmetric tensor **E**. The lattices $\{\mathbf{X}\}$ and $\{\mathbf{X}'\}$ differ only by their orientation relative to the lattice $\{\mathbf{r}\}$. Therefore, taking **E** in Eq. (3) as given,[1)] it is possible to determine a set $\{\{\mathbf{X}(\mathbf{\Omega})\}\}$ of lattices enumerated by the transformation $\mathbf{\Omega}$ and to pose the question of choosing the lattice $\{\mathbf{X}(\mathbf{\Omega}_0)\}$ closest to $\{\mathbf{r}\}$. Actually, this is a question of the prominence of an orientational correspondence between lattices for which the difference between the lattices will be minimum. This question reduces to solving the problem of an extremum (minimum) of the measure $m(\{\mathbf{X}(\mathbf{\Omega})\}, \{\mathbf{r}\}) = m(\mathbf{L}(\mathbf{\Omega}), \mathbf{I})$ as a function of $\mathbf{\Omega}$, where $\mathbf{\Omega}$, being a rotation, operates on an arbitrary vector **a** according to the rule $\mathbf{\Omega}\mathbf{a} = [\mathbf{I} \cos \varphi + (1 - \cos \varphi)\mathbf{l} \cdot \mathbf{l}]\mathbf{a} + \sin \varphi[\mathbf{l}, \mathbf{a}]$, $0 \leq \varphi < 2\pi$ is the rotation angle, and the unit vector **l** determines the orientation of the rotation axis.[2)] The necessary and sufficient condition for a minimum can be expressed in this case by the equation Tr $(\mathbf{E}d\mathbf{\Omega}) = 0$. It is equivalent to the system of equations $(d\mathbf{l}, E\mathbf{l})(1 - \cos \varphi) = 0$, Tr **E** $= (\mathbf{l}, E\mathbf{l})\sin \varphi = 0$ for φ and **l**. The latter system is satisfied identically for $\varphi = 0$ and arbitrary **l**. This solution corresponds to $\mathbf{\Omega}_0 = \mathbf{I}$. Therefore among the lattices $\{\{\mathbf{X}(\mathbf{\Omega})\}\}$ the lattice $\{\mathbf{X}(\mathbf{I})\} = \{\mathbf{X}'\}$ meets the requirement of maximum closeness to the lattice $\{\mathbf{r}\}$ and hence the orientational correspondence between the lattices associated by the deformation **E** itself is prominent.

4. We now turn to the case of fcc and bcc(bct) lattices (or γ and α in more succinct notations). The difference between them will be minimal for the Bain orientational correspondence. A direct check for the lattices α_B, α_N, and α_{K-Z} of three orientations relative to the γ lattice which correspond to the Bain, Nishiyama, and Kurdyumov–Sachs orientational relations (ORs), for example, gives

$$m(\alpha_B, \gamma) < m(\alpha_N, \gamma) < m(\alpha_{K-Z}, \gamma), \tag{4}$$

where $m(\alpha, \gamma) = \{\text{Tr}\,[(\mathbf{L} - \mathbf{I})(\mathbf{L}^* - \mathbf{I})]\}^{1/2}$, $\mathbf{L} = \mathbf{\Omega}\mathbf{E}$ is a tensor that transforms the γ lattice, and **E** and $\mathbf{\Omega}$ are determined by Eqs. (A1) and (A2) in the Appendix.

Thus, on the basis of the closeness of the γ and α lattices the Bain orientational correspondence should be expected between the γ-austenite and α-martensite lattices in iron alloys. This has been observed for a MT in very thin films.[14] However, for α-martensite lattices in bulk samples orientations close to α_N or α_{K-Z} are characteristic. However, the order relation established by the inequalities (4) for the lattices α_B, α_N, and α_{K-Z} was obtained assuming that the γ lattice is perfect. It can change if the γ lattice is imperfect. It is of interest to study such cases, since the modern concepts of the nucleation of martensite, heterogeneous or homogeneous (in the sense of Ref. 1), are compatible with the ide-

alization of perfect γ-austenite as a model of the structural state preceding a MT.

5. Let us assume that the distortions of austenite are due to a screw dislocation with the direction $\gamma=\langle 110\rangle_\gamma/\sqrt{2}$ characteristic for fcc lattices. The distortion tensor χ, describing in the continuous approximation elastic distortions of the lattice around an infinite screw dislocation with the line $\langle 110\rangle_\gamma$ and Burgers vector $\mathbf{b}=b\beta$ in an infinite crystal with cubic symmetry, is given by the formula $\chi=2\tilde{\varepsilon}\beta\cdot\mathbf{i}$, where $\mathbf{i}=[\rho,\tau]$, $\rho=\mathbf{r}/r$, $\mathbf{r}=[\tau,[\mathbf{r}_O,\tau]]$, the radius vector $\mathbf{r}_O$ determines the position of a point in perfect austenite relative to the pole O chosen on the dislocation line, $\tilde{\varepsilon}=bA^{1/2}(4\pi rf)$, A is the elastic anisotropy parameter,

$$f=A+2(1-A)\sum_{n=1}^{3}(\mathbf{e}_n,\tau)^2(\mathbf{e}_n,\mathbf{i})^2, \tag{5}$$

and $\{\mathbf{e}_n\}$ is a right-handed triplet of unit vectors directed along the four-fold symmetry axes.

We denote by γ' the lattice of imperfect austenite, and we shall treat it as being the result of the mapping $\gamma\rightarrow\gamma'$ of the lattice of perfect austenite, described by the tensor $\mathbf{T}=\mathbf{I}+\chi$. The location of points near which the measure of the difference between the lattices $\gamma'(\mathbf{T})$ and $\alpha(\mathbf{L})$ with given orientation reaches an absolute minimum is determined by solving the equation Tr $\{(\mathbf{L}^*-\mathbf{T})[(\partial\mathbf{T}/\partial r)dr+(d\rho,\partial/\partial\rho)\mathbf{T}]\}=0$, expressing the necessary and sufficient condition for a minimum of the measure $m(\alpha(\mathbf{L}),\gamma'(\mathbf{T}))=\mathrm{Tr}\{[(\mathbf{L}-\mathbf{T})(\mathbf{L}^*-\mathbf{T}^*)]\}^{1/2}$. This equation is equivalent to the system of equations $(\rho,(\mathbf{L}^*-\chi^*)\beta)=0$, $(\mathbf{i},(\mathbf{L}^*-\chi^*)\beta)=0$ for r and Ω, satisfied identically at $r=r_a$ and $\rho=\rho_a$, where $r_a=2b\tilde{\varepsilon}_a/J$, $\rho_a=[\tau,\mathbf{i}_a]$, $\mathbf{i}_a=(\beta,\tau)\mathbf{J}/J$, $\mathbf{J}=[\tau,[\mathbf{L}^*\tau,\tau]]$, $\tilde{\varepsilon}_a=A^{1/2}/(4\pi f_a)$, and f_a is the value of the function (5) at $\mathbf{i}=\mathbf{i}_a$. Elastic lattice distortions of perfect austenite at points located at a distance r_a in the direction ρ_a from the dislocation line are described by the tensor $\mathbf{T}=\mathbf{T}_a$, where $\mathbf{T}_a=\mathbf{I}+\tau\cdot\mathbf{J}$. The tensor $\mathbf{T}=T_a$ maps the lattice γ into the lattice $\gamma'_a=\gamma'(\mathbf{T}_a)$, which is maximally close to the lattice $\alpha(\mathbf{L})$. The difference between them is characterized by the smallest value of the measure $m(\alpha(\mathbf{L}),\gamma'(\mathbf{T}))$

$$m_a(\alpha(\mathbf{L}),\gamma'_a)=[m^2(\alpha(\mathbf{L}),\gamma)-J^2]^{1/2}. \tag{6}$$

For given $\mathbf{L}=\Omega\mathbf{E}$ (i.e., for a given relative orientation of α and γ lattices in this case) not all dislocations from the family under consideration bring into coincidence the lattices $\alpha(\mathbf{L})$ and γ in the same way at the indicated points. The smallest of the values (6) (which we denote by $m_a(\alpha_B,\gamma'_a)$, $m_a(\alpha_N,\gamma'_a)$, and $m_a(\alpha_{K-Z},\gamma'_a)$) corresponds to dislocations (four, two, and one respectively) whose lines make the smallest angles with the directions appearing in the Bain, Nishiyama, and Kurdyumov–Sachs orientational relations. However, calculations using the data[5] on the parameters of γ and α lattices show that $m_a(\alpha_{K-Z},\gamma'_a)<m_a(\alpha_N,\gamma'_a)$, $m_a(\alpha_B,\gamma'_a)$.

Thus the order relation for the α_B, α_N, and α_{K-Z} lattices is reversed if they are compared with the lattice γ'_a and the Kurdyumov–Sachs orientation becomes prominent provided that the direction Λ appearing in the Kurdyumov–

Sachs order relation is parallel to the dislocation line. We shall confine our attention to this case, taking for definiteness $\mathbf{p}=[001]_\gamma$, $\Lambda=-\tau=[\bar{1}01]_\gamma/\sqrt{2}$, and $\mathbf{N}=[111]_\gamma/\sqrt{3}$ in Eqs. (A1) and (A2). The distance r_a and the direction ρ_a will then be given by

$$r_a=bA^{1/2}(2\pi J), \quad \rho_a=(\beta,\tau)[0\bar{1}0]_\gamma, \tag{7}$$

where $J=\kappa(2-t^2)[2(2+t^2)]^{-1/2}$. The values of J lie in the range 0.27–0.33, depending on the values of κ and t, and r_a/b lies in the range $(0.48-0.58)A^{1/2}$. We note that the dislocation-induced distortions of austenite are nonuniform and decrease with distance from the dislocation line. Therefore the region where the Kurdyumov–Sachs correspondence between the lattices α and γ' has not yet lost out to the Bain correspondence should befinite in the transverse section $(\bar{1}01)_\gamma$. It is convenient to give the positions of the points in this section by the radius vectors $\mathbf{r}=r(\rho_a\cos\psi+\mathbf{i}_a\sin\psi)$ in the basis formed by the vectors ρ_a [see Eq. (7)] and $\mathbf{i}_a=(\beta,\tau)[101]_\gamma/\sqrt{2}$, where r and ψ are determined by the condition $m(\alpha_{K-Z},\gamma'(\mathbf{T}))\leqslant m(\alpha_B,\gamma)$, which in turn leads to the inequalities

$$\Gamma_-(\psi)\leqslant r/r_a\leqslant\Gamma_+(\psi), \quad -\psi_c\leqslant\psi\leqslant\psi_c,$$

where

$$\Gamma_\pm=[\cos\psi\pm(\cos^2\psi-\cos^2\psi_c)^{1/2}]/(f\cos^2\psi_c),$$

$$\cos^2\psi_c=[m^2(\alpha_{K-Z},\gamma)-m^2(\alpha_B,\gamma)]/J^2,$$

and $f=A-(A-1)\cos^2\psi$.

The angle ψ_c as a function of κ and t varies over the range $(0.190-0.195)\pi$. The maximum size of the section under study in the direction ρ_a is given by the expression $2r_a\sin\psi_c/\cos^2\psi_c$ and lies in the range $(1.64-1.73)r_a$. The maximum size in the direction $\mathbf{i}_a$ can be estimated from the expression

$$(2/A)^{1/2}r_a\{2\sin^2\psi_c-A^{-1}$$
$$+[A^{-2}+4(1-A^{-1})\sin^2\psi_c]^{1/2}\}^{1/2}/\cos\psi_c.$$

For $A\gg1$ this size is essentially independent of A and is approximately $1.5b$.

Thus comparing the α_B, α_N, and α_{K-Z} lattices with the austenite lattice, which differs from a perfect lattice near a screw dislocation, shows that the Kurdyumov–Sachs orientation is prominent in a cylindrical (noncircular) region oriented parallel to the dislocation line. In the transverse section this region is characterized by two dimensions: the smallest dimension in the direction $\mathbf{i}_a$ and the largest in the direction ρ_a. The first dimension is essentially independent of the elastic anisotropy parameter and is close to $1.5b$. The second one is proportional to $A^{1/2}$ and reaches 20 to 30 times b only for $A\geqslant300$. Therefore these dimensions are comparable to those which are usually taken[16] for α-martensite nuclei, if the elastic anisotropy is large. Such an anisotropy is not characteristic for the volume elastic properties of iron alloys [for example, in Fe–Ni Invar alloys $A\approx4$ (Ref. 17)], but in principle it is realizable in systems with quite appreciable softening of the elastic modulus C' near the MT onset temperature. The existence of local regions near defects where partial softening of the elastic moduli and a decrease of stability can be expected[18] also has not been ruled out.

If coincidence of the α and γ lattices decreases the nucleation barrier for α-martensite, then the elastic field of a screw dislocation gives the most favorable conditions for nucleation of martensite with the α_{K-Z} lattice, giving rise not only to deformation but also rotation of the of the austenite lattice. The deformation is described by the strain tensor $\boldsymbol{\varepsilon} = (\boldsymbol{\chi} + \boldsymbol{\chi}^*)/2$ and consists of stretching $\varepsilon_1 = \tilde{\varepsilon}$ and compression $\varepsilon_2 = -\tilde{\varepsilon}$ in the directions $\boldsymbol{\xi}_1 = (\mathbf{i} + \boldsymbol{\beta})/\sqrt{2}$ and $\boldsymbol{\xi}_2 = (\boldsymbol{\beta}, \boldsymbol{\tau})(\mathbf{i} - \boldsymbol{\beta})/\sqrt{2}$. For a dislocation with the line $\boldsymbol{\tau} = [10\bar{1}]_\gamma/\sqrt{2}$, specifically, $\tilde{\varepsilon} = bA^{1/2}\{4\pi r[A + (1 - A) \times \cos^2\psi]\}^{-1}$, $\mathbf{i} = \mathbf{i}_a\cos\psi - \boldsymbol{\rho}_a\sin\psi$, where $\mathbf{i}_a$ and $\boldsymbol{\rho}_a$ are defined above. Therefore for fixed r and $A > 1$ stretching and compression in the direction $\boldsymbol{\rho} = \boldsymbol{\rho}_a$, where $\boldsymbol{\rho} = \boldsymbol{\rho}_a\cos\psi + \mathbf{i}_a\sin\psi$, reach maximum values, while $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ coincide (if $\boldsymbol{\beta} = \boldsymbol{\tau}$) with the directions $[100]_\gamma$ and $[001]_\gamma$ characteristic for stretching and compression under a Bain deformation with the compression axis $\mathbf{p} = [001]_\gamma$. This agrees with the directions $\boldsymbol{\Lambda} = [\bar{1}01]_\gamma/\sqrt{2}$ and $\mathbf{N} = [111]_\gamma/\sqrt{3}$ in the Kurdyumov–Sachs orientational relation. The latter is important from the standpoint of the elastic model of a dislocation center of nucleation (DCN) of martensite.[3] In the DCN model the locations favorable for nucleation are associated with the neighborhood of points where stretching and compression of austenite occur in orthogonal directions close to the extremal directions for the given dislocation. Therefore, for a screw dislocation the results concerning favorable locations for nucleation of α-martensite that follow from the DCN model are consistent with the results obtained on the basis of the definition (2) of the measure of the closeness of lattices. when comparing with the elastic model of DCN it should also be remembered that for the orientational variants of an α lattice which were considered above the orientations for which the direction appearing in the orientational relation makes the smallest angle with the dislocation line are always prominent. This independently confirms the results of Ref. 19 with respect to a screw dislocation as well as the ideas of Ref. 3 concerning the relationship between the elastic state of a nucleation center and the morphology of martensite.

6. We shall now summarize. The investigation of the relative orientation of the lattices of α-martensite and γ-austenite, which are directly related with their closeness, leads to reasonable results for perfect and imperfect lattices of γ-austenite. This shows that the definition (2) of the measure of closeness of lattices is fruitful.

A concept of the closeness of lattices which is expressed in a quantitative form expands the possibilities of a geometric approach to studying the effect of the elastic field of a defect on the barrier for martensite nucleation, making it possible to establish an order relation for the lattices of imperfect austenite and a given lattice. In addition, the latter may not coincide with the martensite lattice and is postulated on the basis of ideas about deformations giving rise to the most favorable conditions for a transition to a new structural state at the nucleation stage, as is done, for example, in the DCN model. The deformation postulated in the DCN model differs substantially from the Bain deformation both with respect to the character (deformation with weakly distorted planes) and magnitude $\sim 10^{-4} - 10^{-3}$ of the threshold

stretching and compression. The use of the concept of closeness of lattices in application to the DCN model and also for investigating the most favorable variants of partially coherent conjunction of the α-martensite and γ-austenite lattices can serve as a subject of a separate analysis.

APPENDIX

The tensor $\mathbf{E}$ describing the $\gamma \rightarrow \alpha$ restructuring according to Bain can be expressed by the formula

$$\mathbf{E} = \kappa\sqrt{2}\{\mathbf{I} + [(t/\sqrt{2}) - 1]\mathbf{p} \cdot \mathbf{p}\}, \tag{A1}$$

where $\kappa = a_\alpha/a_\gamma$, $t = c_\alpha/a_\alpha$, a_γ is the parameter of the γ lattice, a_α and c_α are parameters of the α lattice, and $\mathbf{p} = \langle 100 \rangle_\gamma$ gives the orientation of the axis of Bain compression.

The rotation $\boldsymbol{\Omega}$ that gives after a Bain deformation the Nishiyama and Kurdyumov–Sachs orientational correspondence between the γ and α lattices is given by the formula

$$\boldsymbol{\Omega} = \mathbf{N} \cdot \mathbf{e} + \boldsymbol{\Lambda} \cdot \mathbf{e}' + [\mathbf{N}, \boldsymbol{\Lambda}] \cdot [\mathbf{e}, \mathbf{e}'], \tag{A2}$$

where $\mathbf{e} = (\mathbf{N}, \mathbf{E}^{-2}\mathbf{N})^{-1/2}\mathbf{E}^{-1}\mathbf{N}$; $\mathbf{e}' = (\boldsymbol{\Lambda}, \mathbf{E}^2\boldsymbol{\Lambda})^{-1/2}\mathbf{E}\boldsymbol{\Lambda}$; $\mathbf{N} = \langle 111 \rangle_\gamma/\sqrt{3}$ is a vector normal to the plane appearing in the Nishiyama and Kurdyumov–Sachs orientational relation; $\boldsymbol{\Lambda}$ gives the direction appearing in the orientational relation, $\boldsymbol{\Lambda} = \sqrt{6}[\mathbf{N}, [\mathbf{p}, \mathbf{N}]]/2$ for the Nishiyama orientational correspondence and $\boldsymbol{\Lambda} = \sqrt{6}\{[\mathbf{p}, \mathbf{N}] \pm \sqrt{3}[\mathbf{N}, [\mathbf{p}, \mathbf{N}]]\}/4$ for the Kurdyumov–Sachs correspondence.

---

[1]See Ref. 13 concerning approaches to the crystallography of martensite and the construction of the tensors $\mathbf{E}$.

[2]The symbols $\mathbf{a} \cdot \mathbf{b}$, $(\mathbf{a}, \mathbf{b})$, and $[\mathbf{a}, \mathbf{b}]$ employed in the text denote tensor, scalar, and vector products, respectively. The action of the tensor $\mathbf{a} \cdot \mathbf{b}$ on the vector $\mathbf{c}$ is determined by the rule $\mathbf{a} \cdot \mathbf{b}\mathbf{c} = \mathbf{a}(\mathbf{b}, \mathbf{c})$.

[1] V. V. Kondrat'ev and V. G. Pushin, Fiz. Met. Metalloved. **60**, 629 (1985).
[2] V. P. Vereshchagin and S. M. Kashchenko, Fiz. Tverd. Tela (Leningrad) **31**, 287 (1989) [Sov. Phys. Solid State **31**, 885 (1989)].
[3] V. P. Vereshchagin and M. P. Kashchenko, Metalloved. Term. Obrab. Met., **7**, 5–11 (1994).
[4] V. I. Sarrak and S. O. Suvorova, Izv. Akad. Nauk. SSSR, Met. **6**, 90 (1982).
[5] A. A. Leont'ev, V. M. Schastlivtsev, and L. N. Romashov, Fiz. Met. Metalloved. **58**, 950 (1984).
[6] M. P. Kashchenko, Fiz. Met. Metalloved. **58**, 852 (1984).
[7] K. E. Easterling and A. R. Thölen, Acta Metall. **24**, 333 (1976).
[8] E. Z. Vintaĭkin, in *Progress in Science and Technology. Metallurgy and Heat Treatment*, VINITI, Moscow, 1983, Vol. 17, pp. 3–63.
[9] K. E. Easterling and H. M. Miekk-oja, Acta Metall. **15**, 1133 (1967).
[10] R. Monzen, A. Sato, and T. Mori, Trans. Jpn. Inst. Met. **22**, 65 (1981).
[11] V. P. Vereshchagin, Izv. Vyssh. Uchebn. Zaved. Fiz. (Tomsk), 1996; deposited in VINITI, No. 3010-V96, Moscow, 1996, 16 pp.
[12] C. Teodosiu, *Elastic Models of Crystal Defects* [Springer-Verlag, New York (1982); Mir, Moscow (1985), 352 pp.].
[13] J. W. Christian, in *Physical Metallurgy* [Mir, Moscow, 1968, No. 2, pp. 227–346].
[14] Z. Nishiyama, K. Shimizu, and K. Sugino, Acta Metall. **9**, 620 (1961).
[15] N. J. Gu, in *Proceedings of the International Conference on Martensitic Transformations,* Institute of Metals, Japan, 1986, pp. 325–330.
[16] Yu. N. Petrov, *Defects and Diffusion-Free Transformation in Steel*, Naukova Dumka, Kiev, 1978, 262 pp.
[17] G. Hausch and H. Warlimont, Acta Metall. **21**, 401 (1973).
[18] G. Guenin and P. C. Clapp, in *Proceedings of the International Conference on Martensitic Transformations,* Nara, Japan, 1986, pp. 171–179.
[19] V. P. Vereshchagin and M. P. Kashchenko, Fiz. Tverd. Tela (Leningrad) **33**, 1605 (1991) [Sov. Phys. Solid State **33**, 906 (1991)].

Translated by M. E. Alferieff

Electroluminescence kinetics of film structures based on manganese-doped zinc sulfide

N. T. Gurin and O. Yu. Sabitov

Ul'yanovsk State University, 432700 Ul'yanovsk, Russia

A theoretical and experimental study is made of the growth and decay constants of the brightness as functions of the rise time and amplitude of the pulse from electroluminescent structures based on manganese-doped zinc sulfide and deposited on smooth and rough substrates, and as functions of the rise time and amplitude of the ramped (linearly rising) exciting voltage. The curves obtained are used to determine a variety of parameters and characteristics of the electroluminescence process: the lifetime of the excited luminescence centers, the excitation and relaxation probabilities of luminescence centers per unit time, and the cross section of impact excitation of luminescence centers and their dependences on the rise time and amplitude of the linearly rising exciting voltage. Explanations are given for the fact that the indicated characteristics show different behavior for structures on smooth and rough substrates. © *1999 American Institute of Physics.* [S1063-7842(99)01105-8]

To increase the efficiency indicators (brightness, light output, internal and external quantum yields, and energy yield) of film electroluminescent (EL) emitters (ELEs) it is necessary to know the most important parameters determining the kinetics of the electroluminescence, such as the probabilities of excitation and transition into an unexcited state of luminescence centers, the impact excitation cross section of these centers, and others.[1] As shown in Ref. 2, most of these parameters can be determined by exciting film ELEs with a linearly rising voltage. However, since these parameters depend on the excitation regime (shape, amplitude, and temporal parameters of the voltage pulses), a complete description of the electroluminescence kinetics requires knowledge of the indicated dependences. Therefore our objective in the present work was to investigate excitation and relaxation processes for luminescence centers in ZnS:Mn-based film ELEs as a function of the parameters of the linearly increasing exciting voltage for EL structures deposited on ordinary smooth glass substrates and on structures deposited on substrates with a rough surface and having higher electroluminescence efficiency indicators.[3–5]

To solve this problem we performed experimental investigations of metal–insulator–semiconductor–insulator–metal (MISIM) and metal–insulator–semiconductor–composite liquid insulator–metal (MISLM) structures, where M is the first, 0.2 μm thick, SnO$_2$-based transparent electrode deposited on a glass substrate and the second, opaque, 0.15 μm thick Al-based opaque thin-film electrode (for MISLM structures) and a clamped metal electrode with micrometric regulation to within ±5 μm; S–El is a 0.80–0.85 μm thick layer of ZnS:Mn; I — 0.2–0.3 μm thick layer of the insulator ZrO$_2\times$ Y$_2$O$_3$ (13 mass %); L — a 15 μm thick layer of a composite liquid insulator (CLI) consisting of a mixture of PFMS-4 organosilicon liquid with a BaTiO$_3$ powder fill with 1.5–3.0 μm grain size and filled density in the insulator $\sim$50 vol %. The structures were deposited on a smooth substrate and on a substrate with an inner rough surface and an

outer smooth surface. The SnO$_2$ transparent electrode was obtained by hydrolysis of tin chloride. A phosphor layer was deposited by vacuum thermal evaporation in a quasiclosed volume; the opaque electrode was produced by vacuum thermal evaporation; the thin-film insulator layers were produced by electron-beam evaporation; and, the CLI was deposited in the form of a paste. The rough surfaces were obtained by chemical etching of the smooth substrate in hydrofluoric acid. All identical layers of the experimental structures were obtained in the same technological sequence. Measurements with an MII-4 microscope and an FOM-2 photoelectric ocular micrometer showed that the rough substrates possessed on their inner surface 0.2–0.5 μm microasperities, uniformly distributed over the surface, with linear sizes of 0.6–1.0 μm; the distribution of the height and linear dimensions of the microasperities corresponded to a Gaussian distribution. The diameter of the opaque electrodes was 2.5 mm. The surface resistance of the transparent electrode was 250 $\Omega/\square$ on substrates with a smooth inner surface and 400 $\Omega/\square$ with a rough surface. Brightness (luminance) was measured witha YaRM-3 luminance–luxmeter with a measurement error of 8%. The parameters of the brightness waves with the EL structures excited by a linearly rising voltage were measured with an FÉU-75 photomultiplier, the signal from which was fed into an S1-114 two-trace oscilloscope. A linearly rising voltage regime was obtained using a symmetric sign-changing trapezoidal voltage with simultaneously adjustable equal rise and fall times of the pulse[2,6] applied from a G5-89 generator with an additional amplifier. The measurement error of the temporal parameters was $\pm5\%$. An MUM-2 monochromator with 0.6 nm resolution and ±0.5 nm measurement error was used to measure the spectral characteristics.

To determine the dependences of the basic electroluminescence parameters on the parameters of the exciting voltage we shall examine the kinetics of electroluminescence in

film ELEs excited by a linearly rising voltage.

Under direct impact excitation of Mn^{2+} ions the change in the density of excited luminescence centers is described by the equation[2]

$$\frac{dN^*(t)}{dt} = \alpha(t)[N - N^*(t)] - \frac{N^*(t)}{\tau} - \beta N^*(t), \qquad (1)$$

where N is the density of luminescence centers, $N^*(t)$ is the density of excited luminescence centers, $\alpha(t) = \sigma j(t)/e$ is the probability of a transition of a luminescence center from the ground into an excited state per unit time (σ is the impact excitation cross section of a luminescence center, $j(t)$ is the conduction current density in the phosphor film, giving rise to luminescence of the film, and e is the electron charge); τ is the relaxation time constant of the excited luminescence centers which is due to the radiative transitions into the ground state; and, β is the probability of nonradiative transitions of these centers per unit time.

For the analysis, we assume that the insulators are uniform and have no dielectric losses, the electroluminescent structure is symmetric, and the brightness waves are identical in different half periods of the exciting voltage.

We shall solve Eq. (1), similarly to Ref. 2, under the assumption that σ does not depend on t. Since the lifetime τ^* of the excited luminescence centers is

$$\tau^* = \left(\frac{1}{\tau} + \beta\right)^{-1}, \qquad (2)$$

and using the expression $j(t) = C_{d0}(dV/dT) = C_{d0}(V_m/t_m)$ for the active current density on the linearly rising section of the exciting voltage,[2,6] where C_{d0} is the specific capacitance of the insulator layers of the EL structure and t_m is the time at which the excitation voltage reaches an amplitude V_m (the moment of the transition to the flat top of the trapezoidal pulse), we obtain the density of excited centers

$$N_H^*(t) = C \exp\left[-\left(\alpha + \frac{1}{\tau^*}\right)t\right] + \frac{\alpha N}{\alpha + 1/\tau^*} \qquad (3)$$

on the section where the exciting voltage increases up to its maximum value V_m and

$$N_C^*(t) = C \exp\left(-\frac{t}{\tau^*}\right) \qquad (4)$$

on the section corresponding to the flat top and the dropoff of the voltage pulse. Here C are constants determined by the initial conditions.

We now substitute into Eq. (5) the solutions (3) and (4) taking account of the initial conditions, which take account of the onset of radiation generation when the exciting voltage reaches the threshold value. The equation (5) relates the brightness of the ELE luminescence and the density of excited luminescence centers under monochromatic radiation and constant brightness in all directions (a Lambert source)

$$L(t) = \pi \eta_c d_l \frac{N^*(t)}{\tau^*}, \qquad (5)$$

where η_c is the light output efficiency given by

$$\eta_c = \eta_{ext} f_\lambda h \nu, \qquad (6)$$

η_{ext} is the external quantum yield, equal to the ratio of the number of photons emitted from the surface to the total number of photons arising in the volume of the EL layer; f_λ is the visibility of the radiation; $h\nu$ is the energy of the emitted photons (h is Planck's constant); and, d_l is the thickness of the phosphor layer.

As a result, we obtain the instantaneous brightness as a function of time with the exciting voltage rising from the threshold value V_n, corresponding to the time t_n, when the thin-film ELE starts to luminesce ($L_H(t_n) = 1$ cd/m²), up to the time t_m when the voltage amplitude reaches the value V_m

$$L_H(t) = \frac{\eta_c d_l N}{\tau^*} \frac{\alpha}{\alpha + 1/\tau^*}$$
$$\times \left[1 - \exp\left[-\left(\alpha + \frac{1}{\tau^*}\right)(t - t_n)\right]\right]. \qquad (7)$$

The expression (7), which holds for low repetition frequencies of the exciting voltage pulses, when the brightness drops to virtually zero within a half period, i.e., $(4-5)\tau^* \leqslant T/2$ (T is the repetition period of the exciting voltage pulses), makes it possible to determine the brightness growth constant

$$\tau_g = \left(\alpha + \frac{1}{\tau^*}\right)^{-1}. \qquad (8)$$

Similarly, taking as the initial conditions equal amplitudes of the brightness wave during growth and decay $L_H(t_m) = L_C(t_m) = L_m$ at $t = t_m$, we obtain an expression $L_C(t)$ for the dropoff of brightness with time

$$L_C(t) = \frac{\eta_c d_l N}{\tau^*} \frac{\alpha}{\alpha + 1/\tau^*}$$
$$\times \left[1 - \exp\left[-\left(\alpha + \frac{1}{\tau^*}\right)(t_m - t_n)\right]\right]$$
$$\times \exp\left(-\frac{t - t_m}{\tau^*}\right) = L_m \exp\left(-\frac{t - t_m}{\tau^*}\right). \qquad (9)$$

The lifetime τ^* can be found using expressions (7)–(9).

In general the probability α of a transition of a luminescence center from the ground into an excited state, the impact excitation cross section σ of a center, and the lifetime τ^* of the luminescence centers can depend on the rise time t_m and the amplitude V_m of the exciting voltage. Then the growth constant τ_g of the brightness, its reciprocal $1/\tau_g$, and α can be determined as

$$\tau_g(t_m, V_m) = (\alpha(t_m, V_m) + 1/\tau^*(t_m, V_m))^{-1}, \qquad (10)$$

$$\frac{1}{\tau_g(t_m, V_m)} = \alpha(t_m, V_m) + \frac{1}{\tau^*(t_m, V_m)}, \qquad (11)$$

$$\alpha(t_m, V_m) = \frac{\sigma(t_m, V_m) C_{d0} V_m}{e t_m} = \alpha(t_m, V_m) t_m^{-1}, \qquad (12)$$

whence the expression for determining the impact excitation cross section is

$$\sigma(t_m, V_m) = \left(\frac{1}{\tau_g(t_m, V_m)} - \frac{1}{\tau^*(t_m, V_m)} \right) \frac{e t_m}{C_{d0} V_m}. \quad (13)$$

In accordance with Eqs. (11) and (12) with $\sigma \sim 10^{-16}$ cm^2, $C_{d0} \sim 3 \times 10^{-8}$ F/cm^2, and $V_m \sim 200$ V we have $a(V_m) \approx 4 \times 10^{-3}$ and for $t_m \geq \tau^*$ we have $\alpha(t_m, V_m) \ll 1/\tau^*(t_m, V_m)$. Then the function $1/\tau_g(t_m, V_m)$ reduces to the function $1/\tau^*(t_m, V_m)$

$$\frac{1}{\tau_g(t_m, V_m)} \approx \frac{1}{\tau^*(t_m, V_m)}, \quad \tau_g(t_m, V_m) \approx \tau^*(t_m, V_m). \quad (14)$$

The methodological error does not exceed 0.5%.

Expression (10)–(14) make it possible to find from the experimental dependences $\tau_g(t_m, V_m)$ and $1/\tau_g(t_m, V_m)$ a variety of parameters and characteristics determining the kinetics of electroluminescence in film ELEs. For all structures indicated above, when studying the waves of luminescence brightness of the structures on a sign-changing symmetric trapezoidal voltage with pulse rise and fall times $\tau_m = 20$ μs and pulse repetition frequency 400 Hz, on the saturation section of the brightness versus voltage characteristics (BVCs) the front where the brightness increases to the maximum value was, within the limits of measurement error, exponential, in accordance with Eq. (7), with a growth time constant τ_g. For the MISIM structure on a substrate with a smooth inner surface the decay of the brightness wave after the voltage amplitude V_m was reached was also exponential, in accordance with Eq. (9), with time constant 650 μs. For an MISLM structure on a smooth substrate and for all EL structures on substrates with a rough inner surface the decay of the brightness wave possessed two exponential sections: a fast section with decay constant $\tau_{c1} = 200$–260 μs and a slow section with $\tau_2 = 460$ μs (MISIM structure), 400 μm (MISLM structure on a smooth substrate), and 300 μs (MISLM structure on a rough substrate).

The form of the experimental dependences $\tau_g(t_m)$, measured as a function of t_m in the range 1–300 μs with a 400 Hz repetition frequency of these pulses (Fig. 1), attests to a large increase of τ_g with increasing t_m, especially for the MISIM structure on a smooth substrate. The saturation sections of the dependence $\tau_g(t_m)$ for large t_m (300 μs) make it possible to determine the lifetimes τ^* of the excited luminescence centers in accordance with Eq. (14). For an MISIM structure on a smooth substrate (Fig. 1a) τ^* is 0.9 ms on the increasing section of the BVC ($V_m = 90$ V) and 0.84 ms on the saturation section ($V_m = 140$ V). The reciprocals ($1/\tau^*$) determine the probability of transitions of luminescence centers from an excited into an unexcited state per unit time and are 1110 and 1190 s^{-1} on the rising and saturation sections of the BVC, respectively. The methodological error in τ^* and $1/\tau^*$, taking account of the values presented for t_m and τ^*, does not exceed 1.5%. For a rough surface the values of τ_g decrease and the saturation section of $\tau_g(t_m)$ shifts in the direction of large values of t_m; this could be due to both a decrease of τ^* and an increase of α, for example, as a result
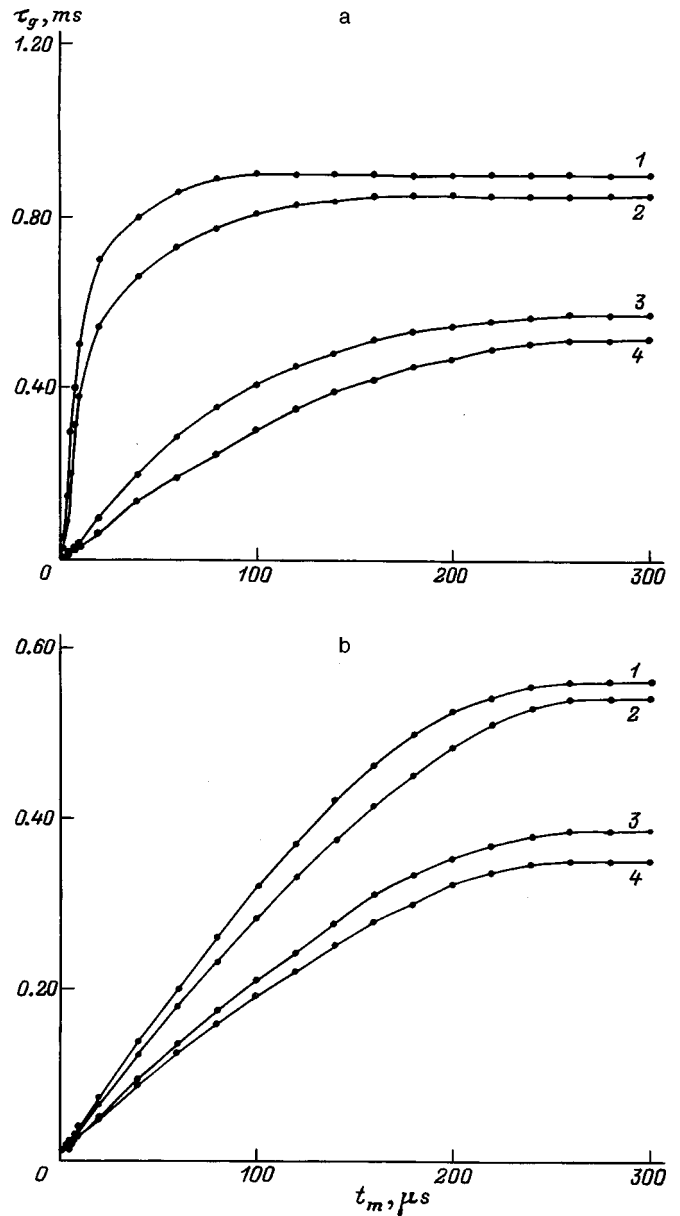

FIG. 1. τ_g versus τ_m: a — for an MISIM structure; b — for an MISLM structure; *1,3* — on the rising section of the BVC; *2,4* — on the saturation section of the BVC; *1,2* — on a smooth substrate; *3,4* — on a rough substrate.

of an increase in the effective specific capacitance C_{d0} of the insulator layers and also an increase of σ. Similar behavior is also observed for the MISLM structure on smooth and rough substrates (Fig. 1b), the latter also leading to an additional decrease of τ_g. For the MISIM structure on a substrate with a rough inner surface (Fig. 1a) $\tau^* = 0.57$ ms ($1/\tau^* \approx 1750$ s^{-1}) on the rising section of the BVC and $\tau^* = 0.51$ ms ($1/\tau^* \approx 1960$ s^{-1}) on the saturation section; for the MISLM structure (Fig. 1b) on a smooth substrate $\tau^* = 0.56$ ms ($1/\tau^* \approx 1790$ s^{-1}) on the increasing section of the BVC and $\tau^* = 0.53$ ms ($1/\tau^* \approx 1890$ ms^{-1}) on the saturation section; on a rough substrate $\tau^* = 0.38$ ms ($1/\tau^* \approx 2630$ ms^{-1}) on the increasing section of the BVC and $\tau^* = 0.35$ ms ($1/\tau^* \approx 2860$ ms^{-1}) on the saturation section. The measurements performed with the E7-14 apparatus attest to the fact that the

capacitance of MISIM and MISLM structures on rough substrates is 20–30% higher than for structures on smooth substrates. This confirms that the increase in C_{d0} plays a role in the increase of α.

As follows from the data presented above and in Fig. 1, for large values of t_m, where the relation (14) holds, the lifetime τ^* of the excited luminescence centers of all structures depends on the voltage V_m and is independent of t_m. For this reason the general decrease in the values of τ^* for EL structures on a rough substrate as compared with structures on smooth substrates (Fig. 1) can most likely be explained by an increase of the electric field in the EL layer at the locations of microasperities on the substrate.[3–5] The lower values of τ^* for the MISLM structure on a smooth substrate as compared with the MISIM structure on the same substrate are explained by the presence of a nonuniform electric field in such a MISLM structure because of the nonuniform distribution of the permittivity of the CLI layer.[7,8]

The possible mechanisms leading to a decrease of the brightness growth constant τ_g on switching from smooth to rough substrates can be distinguished most easily, on the basis of Eqs. (11) and (12), according to the experimental curves $1/\tau_H(t/t_m)$. These curves $1/\tau_g(1/t_m)$ (Fig. 2), in complete agreement with Eqs. (11) and (12), are linear for an MISIM structure on a smooth substrate in the entire range $\tau_m \leqslant 300$ μs, for an MISLM structure on a smooth substrate for $t_m \leqslant 50$ μs, and on a rough substrate for $t_m \leqslant 100$ μs. For the MISIM structure the slope of the curve $1/\tau_g(1/t_m)$ increases by a factor of ~ 1.4 with the voltage V_m increasing by a factor of ~ 1.55. Taking account of the errors in the measurements of V_m and τ_g this confirms the validity of expressions (11) and (12) for this structure. For the MISLM structures on smooth and rough substrates the slope of the linear sections of the curves $1/\tau_g(1/\tau_m)$ increases by only a factor of ~ 1.1 with V_m increasing by a factor of ~ 1.47. This deviation of the indicated dependences from the working equations (11) and (12) is due to the existence of a dependence $C_{d0}(V_m)$ due to the fact that the permittivity of the CLI fill — barium titanate — decreases with increasing V_m. Specifically, as V_m increases from 1 to 60 V, the specific capacitance, determined experimentally with the E7-14 apparatus, of the CLI layer decreased by a factor of ~ 2.5, giving rise to a corresponding decrease of the resulting value of C_{d0} of the MISLM structure. The data presented above show that τ^* and $1/\tau^*$ for these structures are independent in the indicated range of t_m. For the MISIM structure on a rough substrate, in the range $t_m \leqslant 300$ μs the function $1/\tau_g(1/t_m)$ has the form $1/\tau_g \sim (1/t_m)^2$ on the rising section of the BVC and $1/\tau_g \sim (1/t_m)^3$ on the saturation section. This form of the functions $1/\tau_g(1/t_m)$ attests to dependences of σ on t_m or τ^* on t_m for this structure in the indicated range of t_m.

For the MISLM structure the dependence $1/\tau_g(1/\tau_m)$ becomes weaker for large values of t_m ($t_m \geqslant 50$ μs on a smooth substrate and $t_m > 100$ μs on a rough substrate) (Fig. 2). This is probably due to the appearance of the two decreasing sections, mentioned earlier, of the brightness with lower values of the decay constants τ_{c1} and τ_{c2}, which could be caused by relaxation of, together with single centers, Mn^{2+} pair lumi-
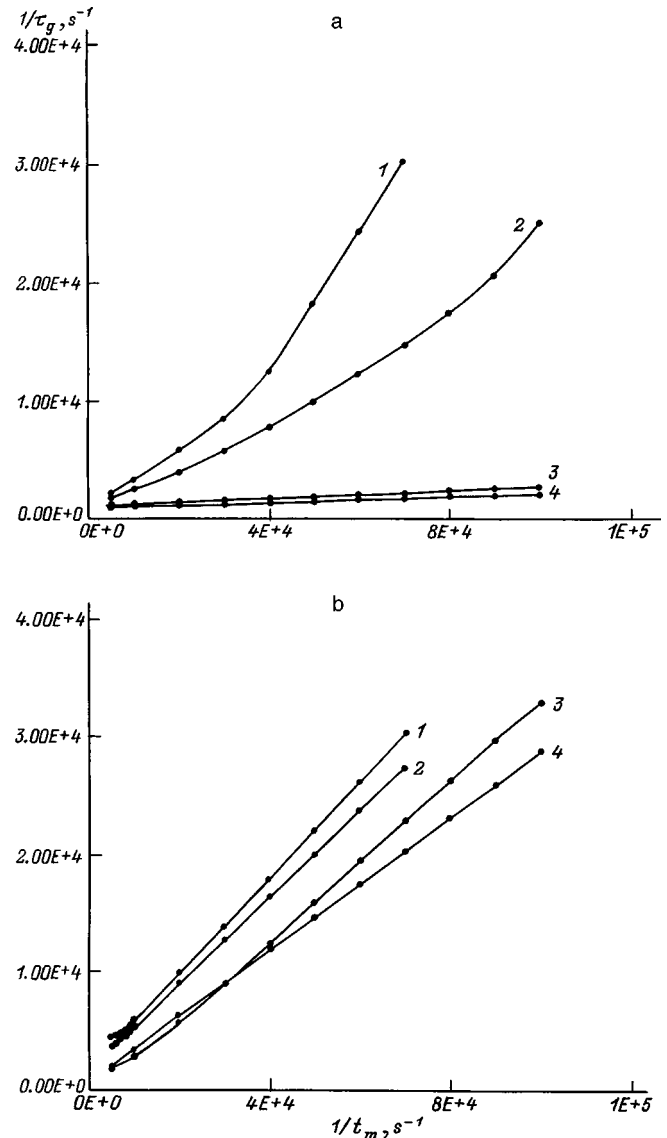

FIG. 2. $1/\tau_g$ versus $1/t_m$: a — for an MISIM structure; b — for an MISLM structure; *1,2* — on a rough substrate; *3,4* —on a smooth substrate; *1,3* — on the saturation section of the BVC; *2,4* — on the rising section of the BVC.

nescence centers[9] with lifetime τ_{c1}; such pair centers could appear as a result of the nonuniform electric field in the EL layer. These centers could have a similar effect on the dependences $1/\tau_g(1/t_m)$ for MISIM structures on a rough substrate (Fig. 2).

Since for the MISIM structure on a smooth substrate and an MISLM structure on smooth and rough substrates Eqs. (11) and (12) hold in the indicated ranges of t_m, in these ranges the experimental curves $1/\tau_g(1/t_m)$ for these structures (Fig. 2) are at the same time dependences of the excitation probability α of the luminescence centers per unit time on $1/t_m$, taking account of the fact that the values of α on the ordinate in this case are obtained by decreasing the corresponding values of $1/\tau_g$ by the constant values of $1/\tau^*$ presented above for these structures.

In summary, for MISIM and MISLM structures $1/\tau_g$ and the slope of linear sections of the curves $1/\tau_g(1/t_m)$ on both
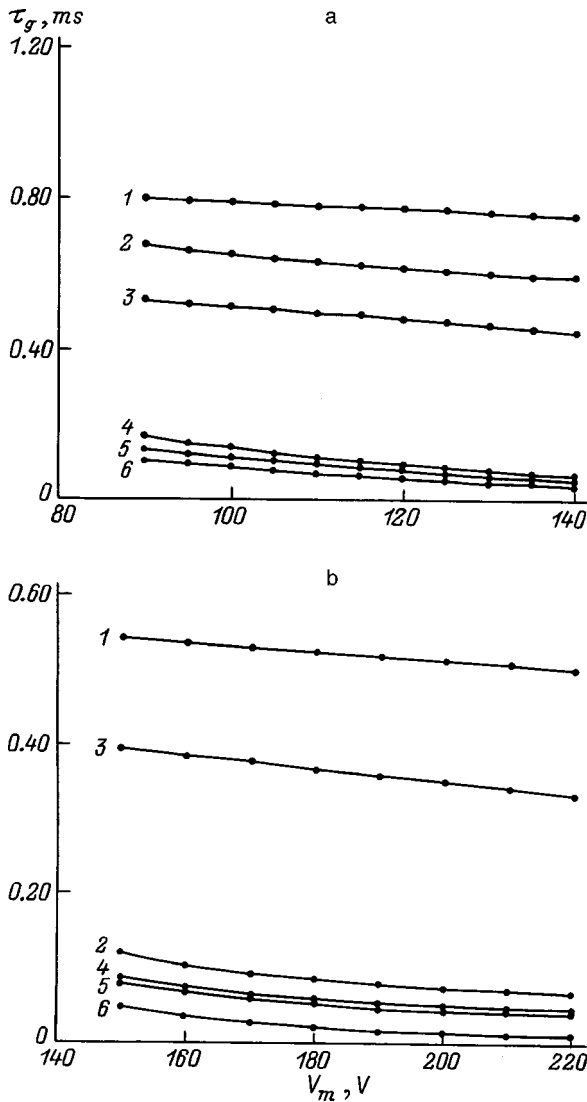
FIG. 3. τ_g versus V_m: a — for an MISIM structure; b — for an MISLM structure; $1,2,4$ — on a smooth substrate; $3,5,6$ — on a rough substrate; $t_m = 300$ $(1,3)$, 20 $(2,5)$, and 1 μs $(4,6)$.
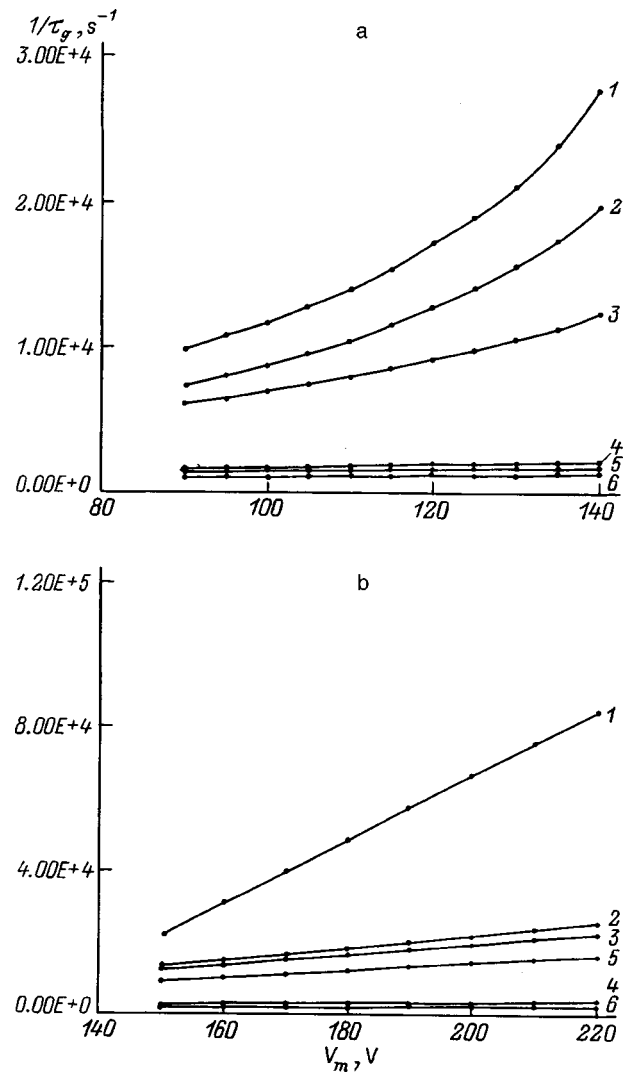

FIG. 4. $1/\tau_g$ versus V_m: a — for an MISIM structure; b — for an MISLM structure; $1,2,4$ — on a rough substrate; $3,5,6$ — on a smooth substrate; $t_m = 1$ $(1,3)$, 20 $(2,5)$, and 300 μs $(4,6)$.

the rising and saturation sections of the BVC tend to increase on switching from a smooth to a rough substrate. This can be explained by an increase of $\alpha(V_m)$ and possibly also of $1/\tau^*(V_m)$ in a nonuniform electric field.

As follows from Eq. (12), for a MISIM structure on a smooth substrate α can increase with V_m on account of V_m itself as well as on account of an increase of $\sigma(V_m)$. To determine the reasons why $\alpha(V_m)$ increases, the dependences of τ_g on V_m were investigated experimentally for MISIM and MISLM structures. According to Eqs. (11) and (12), τ_g decreases with increasing V_m (Fig. 3), and for large t_m (300 μs) these dependences are essentially linear for structures on smooth and rough substrates. Since the relation (14) is satisfied in this case, these dependences are actually V_m dependences of the lifetime τ^* of the excited Mn^{2+} luminescence centers for all EL structures, and the dependences of the reciprocal $1/\tau_g$ on V_m (Fig. 4) with $t_m = 300$ μs are dependences of the transition probabilities $1/\tau^*$ to an unexcited state of the luminescence centers versus V_m for the

corresponding structures. The relations (13) and (14) make it possible to determine the function $\sigma(V_m)$ for an MISIM structure on a smooth substrate, using the relation $\tau_g(V_m) \approx \tau^*(V_m)$ at $t_m = 300\mu$s and the real value $C_{d0} = 3 \times 10^4$ pF/cm^2 and subtracting from the experimental curve $1/\tau_g(V_m)$ measured at $t_m = 1$ μs the experimental curve $1/\tau_g(V_m)$ measured at $t_m = 300$ μs. As indicated above, the methodological error in determining $\sigma(V_m)$ does not exceed 1.5%. The dependence $\sigma(V_m)$ obtained in this manner (Fig. 5) can be approximated by the function $\sigma \approx D(\exp kV_m)/V_m$, where $D \approx 7.22 \times 10^{-15}$ V·cm^{-2} and $k \approx 0.015$ V^{-1}. This form of the function differs from the data in Ref. 1, which gives the function $\sigma \sim \exp(-E_m^{-2})$, where E_m is the electric field in the EL layer. When expression (13) holds for an MISIM structure on a rough substrate, as indicated by the linear character of the experimental curve $1/\tau_g(V_m)$ at $t_m = 300$ μs (Fig. 4), and the capacitance C_{d0} is independent of the voltage V_m, the analogous dependence $\sigma(V_m)$ for this structure (Fig. 5), which can likewise be approximated by the function $\sigma \approx D(\exp kV_m)/V_m$, where $D \approx 5.54 \times 10^{-15}$ V·cm^{-2} and
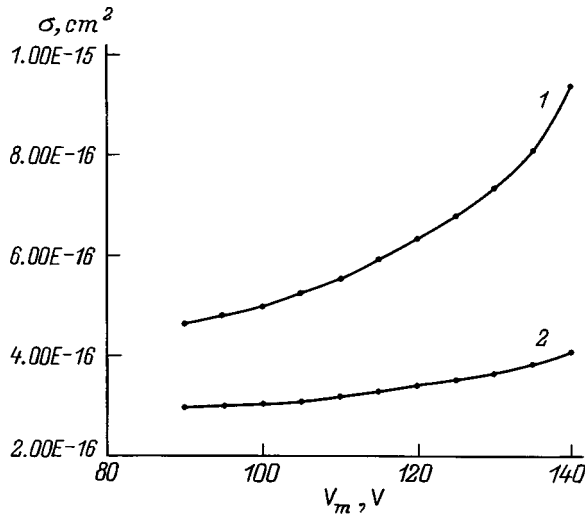
FIG. 5. σ versus V_m: a — for an MISIM structure; b — for an MISLM structure; 1 — on a rough substrate; 2 — on a smooth substrate.

$k \approx 0.022$ V^{-1}, can be estimated. On the whole the impact excitation cross sections σ of Mn^{2+} luminescence centers for an MISIM structure on a rough substrate are 1.6–2.3 times higher than the corresponding values for an MISIM structure on a smooth substrate for the same values of the voltage. The function $\sigma(V_m)$ is more difficult to estimate for an MISLM structure because of the presence of the previously mentioned dependence $C_{d0}(V_m)$.

In contrast to $\sigma(V_m)$ the dependence $\alpha(V_m)$ can be determined in accordance with Eq. (11) for MISIM and MISLM structures on both smooth and rough substrates, since expression (11) in the general form takes account of the excitation probability and the radiative and nonradiative relaxation of the luminescence centers per unit time. Since at $t_m = 300$ μs $1/\tau_g(V_m) \approx 1/\tau_g^*(V_m)$, the difference of the dependences $1/\tau_g(V_m)$ at $t_m = 1 - 300$ μs (Fig. 4), according to Eq. (12), gives $\alpha(V_m)$.

The curves $\alpha(V_m)$ (Fig. 6) obtained by this method from the experimental data (Fig. 4) differ substantially for MISIM and MISLM structures. For an MISIM structure on smooth and rough substrates these dependences can be approximated
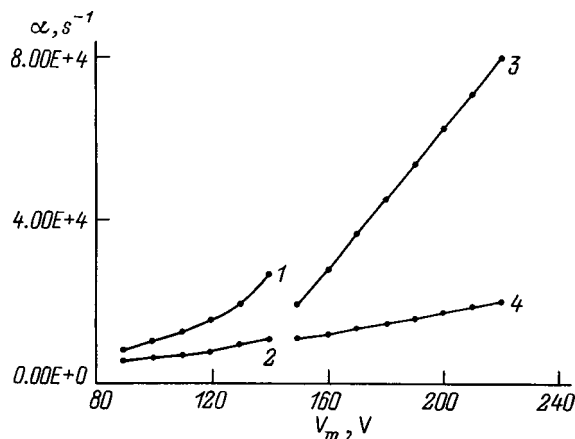

FIG. 6. α versus V_m: $1,2$ — for an MISIM structure; $3,4$ — for an MISLM structure; $1,3$ — on a rough substrate; $2,4$ — on a smooth substrate.

by the function $\alpha \approx \alpha_0 \exp k V_m$, where $\alpha_0 \approx 1353$ s^{-1} and $k \approx 0.015$ V^{-1} for a structure on a smooth substrate and $\alpha_0 \approx 1108$ s^{-1} and $k \approx 0.022$ V^{-1} for a structure on a rough substrate. For the MISLM structure the dependence $\alpha(V_m)$ is linear, $\alpha = k(V_m - V_n) + \alpha_1$. For a structure on a smooth substrate $k \approx 130$ V^{-1}s^{-1} and $\alpha_1 \approx 7 \times 10^3$ s^{-1}; for a structure on a rough substrate $k \approx 870$ V^{-1}s^{-1} and $\alpha_1 \approx 6 \times 10^3$ s^{-1}. The fact that the dependences $\alpha(V_m)$ for an MISLM structure are weaker than for the MISLM structure can be explained by the above-mentioned decrease in the specific capacitance of the CLI layer with increasing V_m, which compensates the increase in $\sigma(V_m)$. The maximum values of α achieved are higher for the MISLM structure than for the MISIM structure. This is explained by the higher intensities of the nonuniform electric field in local sections of MISLM-based ELEs.

It should be noted that the V_m dependences of τ_g, $1/\tau_g$, τ^*, $1/\tau^*$, σ, and α for various values of t_m and also the dependence of $1/\tau_g$ on $1/t_m$ with the corresponding values of V_m are simultaneously dependences on the voltage growth rate V_m/t_m, and they can also be used to describe the kinetics of electroluminescence and to optimize the excitation regime of ELEs.

CONCLUSIONS

Our experimental investigations of brightness waves in MISIM and MISLM structures on smooth and rough substrates with exciting voltages corresponding to the saturation sections of the BVCs attest to an exponential leading edge of the brightness wave in all experimental structures. The decay of the brightness wave for MISIM structures on a smooth substrate is also exponential. Two brightness decay sections with different decay constants are observed in an MISIM structure on a rough substrate; for an MISLM structure there are two decay sections for both smooth and rough substrates. These results can be explained by the presence of a nonuniform electric field in an MISIM structure on a rough substrate and for an MISLM structure on both types of substrates by the formation of Mn^{2+} pair luminescence centers together with single centers.

The theoretical dependences of the brightness growth and decay constants on the rise times and the amplitude V_m of the linearly rising exciting voltage pulse were obtained for ZnS-based El structures with impact excitation of single Mn^{2+} luminescence centers. These dependences are confirmed by the experimental dependences $\tau_g(V_m, t_m)$. The obtained dependences make it possible to determine a variety of parameters characterizing the electroluminescence process: the lifetime of excited luminescence centers, the excitation and relaxation probabilities of the excited luminescence centers per unit time, and the impact excitation cross section of the luminescence centers and their dependences on V_m and t_m.

Specifically, the brightness growth constant increases substantially with the rise time of the exciting voltage pulse and decreases very little with increasing amplitude of the pulse. The lifetime of the excited luminescence centers and the transition probability of the luminescence centers into an

unexcited state per unit time do not depend on t_m in the entire experimental range of t_m ($1-300$ μs) for MISIM structures on a smooth substrate, for $t_m \approx 1-50$ μs for MISLM structures on a smooth substrate, and for $t_m \approx 1-100$ μs for MISLM structures on a rough substrate.

The impact excitation cross sections of luminescence centers for MISIM structures increase with V_m in accordance with the relation $\sigma \approx D(\exp kV_m)/V_m$ for both a structure on a smooth substrate and for a structure on a rough substrate but with different values of the coefficients D and k. The values of σ for a structure on a rough substrate are 1.6–2.3 times higher than the values of σ for a structure on a smooth substrate at the same voltages. For all experimental EL structures the lifetime of the excited luminescence centers decreases slightly according to an essentially linear law; the transition probability of luminescence centers into an unexcited state per unit time increases slightly with V_m, likewise essentially linearly, and these dependences become stronger for structures on rough substrates. The excitation probabilities α of luminescence centers per unit time for MISIM structures depend exponentially on the amplitude V_m of the exciting pulse, while for MISLM structures the dependence is linear, the weakening of the indicated dependence in the latter case being due to a decrease in the specific capacitance of the composite liquid insulator layer with increasing applied voltage. On switching from a smooth to a rough substrate the values of α increase substantially. This is explained by the appearance (for MISIM structure on a rough substrate) and intensification (for the MISLM structure on a rough substrate) of a nonuniform electric field. It follows from the results presented above that the increase in the efficiency (brightness, light output, and so on) of electroluminescence in MISIM and MISLM structures on switching from smooth substrates to substrates with a rough inner sur-

face is due to, besides optical effects giving rise to an increase in the radiation yield from the structure,[4] an increase in the excitation probability α of the luminescence centers per unit time. The latter can be explained both by an increase in the impact excitation cross section σ of these centers and by intensification of the electric field, giving rise to this growth, in the EL layer in local sections corresponding to microasperities of the substrate surface. The probability $1/\tau^*$ of a transition of the luminescence centers from an excited into an unexcited state per unit time also increases but much more weakly than α.

The results obtained can be used to control the parameters of film EL structures and to optimize excitation regimes. The procedure proposed for determining the parameters and characteristics of prebreakdown electroluminescence can be used to investigate similar structures based on other electroluminescent materials.

[1] I. K. Vereshchagin (Ed.), *Electroluminescent Light Sources*, Énergoatomizdat, Moscow, 1990.

[2] N. T. Gurin and O. Yu. Sabitov, Zh. Tekh. Fiz. **69**(2), 58 (1999) [Tech. Phys. **44**, 184 (1999)].

[3] N. T. Gurin, O. Yu. Sabitov, and I. Yu. Brigadnov, Pis'ma Zh. Tekh. Fiz. **23**(15), 7 (1997) [Tech. Phys. Lett. **23**(8), 577 (1997)].

[4] N. T. Gurin and O. Yu. Sabitov, Zh. Prikl. Spektrosk. **64**, 507 (1997).

[5] N. T. Gurin and O. Yu. Sabitov, Pis'ma Zh. Tekh. Fiz. **23**(20), 1 (1997) [Tech. Phys. Lett. **23**, 775 (1997)].

[6] N. T. Gurin and O. Yu. Sabitov, Zh. Tekh. Fiz. **69**(2), 64 (1999) [Tech. Phys. **44**, 190 (1999)].

[7] I. Yu. Brigadnov and N. T. Gurin, Pis'ma Zh. Tekh. Fiz. **16**(23), 71 (1990) [Sov. Tech. Phys. Lett. **16**(12), 914 (1990)].

[8] I. Yu. Grigadnov, N. T. Gurin, and E. B. Ryabinov, Zh. Prikl. Spektrosk. **59**, 175 (1993).

[9] N. A. Vlasenko, Yu. V. Kopytko, and V. S. Pekar, Phys. Status Solidi **81**, 661 (1984).

Translated by M. E. Alferieff

Application of a Gunn-diode current-pulse generator for modulation of semiconductor lasers

S. I. Domrachev, S. A. Alaverdjan, and V. N. Skorokhodov

N. G. Chernyshevski Saratov State University, Scientific and Research Institute of Mechanics and Physics,
410026 Saratov, Russia

Gunn diodes (GDs) are widely used as quite efficient microwave generators in the range from one to hundreds of GHz. The current oscillations of a GD in a nonresonant wide-band circuit consist of a periodic sequence of short pulses whose durations and amplitudes can be tens of picoseconds and several amperes, respectively. The repetition frequency, amplitude, and duration of the pulses depend relatively weakly on the magnitude and character of the load (within certain limits), i.e., a pulsed GD generator is a microwave current generator. Appropriate modifications of the GD and the electrodynamic system make possible external information signal synchronization and control regimes. These properties make it possible to use GDs as fast controlling components in functional devices in microelectronics,[1,2] specifically, for modulation of semiconductor injection lasers (ILs).

Efficient modulation of ILs is difficult because such lasers are a low-resistance nonlinear load with a reactive component. Since the impedance of an IL depends on the signal frequency and amplitude,[3] it is very difficult to phase-match the laser with the pump source in a wide band. The real response time of an IL is also limited by optoelectronic relaxation oscillations,[3] which arise in the laser structure itself or, in the general case, in the IL–control circuit system. Parasitic oscillations can be suppressed or attenuated by using a GD in the current generator mode to modulate the IL. The low impedance of an IL is favorable for stable operation of the GD.

The modulation of an IL in a GD circuit was investigated in Refs. 4–6. In these experiments the current characteristics of the GD and IL were not phase-matched. The lasers were actually controlled by voltage through a passive dividing circuit. Since suboptimal modulation regimes were used, strong relaxation was observed in the output radiation.[6]

In the present paper we investigate the modulation of an IL by short and step current pulses of a Gunn diode in the current generator mode in the control circuit. The results of the modulation of the IL by pulses from a voltage generator are also presented.

Microwave stabilization of the current in the load circuit of a GD is accomplished by rapid restructuring of the electric domains in the GD sample as a result of a change in a circuit parameter: the load resistance, the supply voltage, and others. The restructuring time and, correspondingly, the current

stabilization time do not exceed the domain formation time, which can be estimated as[1]

$$t_f = \frac{3.5}{\mu}\sqrt{L\varepsilon/8\pi qn_0(E_t-E_r)},$$

where L is the diode length, ε is the permittivity of the GD material, μ is the carrier mobility, q is the electron charge, n_0 is the carrier density, E_t is the threshold field (3.2×10^5 V/m in GaAs), and E_r is the residual field in a sample with a domain. For estimation it can be assumed that $E_r \approx E_t/2$.

For the GD used in the present work (the parameters are given below) the formation time (the duration of a current pulse) is $t_f \sim 15$ ps. The transient time increases to t_R.[2] If a resistive load R_L is connected in series with the diode,

$$t_R = t_f\sqrt{1+R_L/R_0},$$

where R_0 is the weak-signal resistance of the GD.

Thus the duration of transient processes is minimal if the load resistance is smaller than the GD resistance. Likewise, the internal resistance of the power source for the GD should be small.

To use the current generator regime efficiently the IL should be connected directly into the GD circuit with no dividers or phase-matching components. The figure of merit of the GD is

$$\eta = \delta I/I_t = (I_t-I_r)/I_t,$$

where I_t is the threshold (maximum) current of the GD and I_r is the residual (minimum) current.

It is theoretically possible to obtain $\eta \approx 0.6$ in a GaAs diode. The experimental values are $\eta = 0.3-0.5$. The quantity δI determines the modulation amplitude of the IL pump current. The maximum degree of modulation of the radiation is achieved when the minimum GD current I_r equals the IL threshold current i_{th}. Then a GD phase-matched with the laser permits pumping the IL with current pulses with amplitude up to $2i_{th}$.

In our experiments we investigated serially produced ILs based on a double heterostructure with a strip contact, specifically, a V-grooved and buried channel type. The typical laser characteristics were as follows: threshold current 40–50 mA, maximum power up to 10 mW at 90–100 mA, and recommended operational power ~ 5 mW at 70 mA. The radiation wavelengths were 0.85 and 1.3 μm. With standard

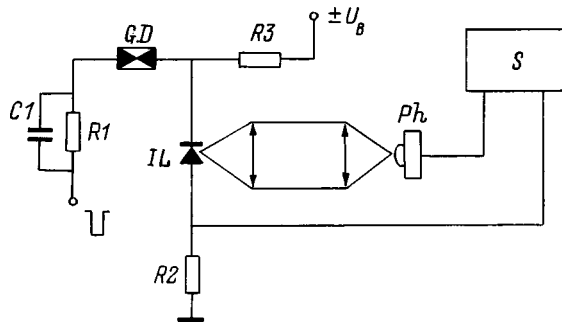FIG. 1. Measurement circuit. Explanations are given in the text.

FIG. 3. Top trace — modulation current (23 mA/div), bottom trace — photodiode response (60 mV/div).

power supplies and pulsed modulation of these lasers self-excited relaxation oscillations with different degrees of sharpness were observed in the radiation.

The modulation scheme investigated (Fig. 1) contains the following components: GD — Gunn diode, IL — injection laser, a charging circuit $C1R1$ for regulating the voltage on the GD, $R2$ — load resistor for monitoring the modulation current, a resistor $R3$ — a circuit for supplying a constant bias voltage to the laser, Ph — photodetector, and S — dual-trace oscilloscope. The electrodynamic system of the mockup of the modulator is formed by a coplanar line, in whose break a controlling GD and a laser diode are connected in series. The system components are arranged on a 22KhS ceramic substrate. An L-2391 pulse generator (a shaper based on a long line with a mercury switch) with a 50 Ω output resistance was used as a source of the bias for the GD–IL circuit. The output radiation of the IL was detected with an LFD-2 avalanche photodiode. Signals from the control resistor $R2$ and from the photodiode load (50 Ω) were observed on an S1-70 sampling oscilloscope.

Phase-matching of the current characteristics of the GD and IL was accomplished by choosing appropriate dimensions (length, cross section) of the crystals used for the GD fabrication. The active resistance of a forward-biased IL was $R_l \approx 5-10 \ \Omega$. The weak-signal resistance of the GD was 200–250 Ω, so that $R_0/R_l \gtrsim 20$. The diodes possessed a planar construction with Ohmic end-face contacts prepared from epitaxial GaAs with density $n_0 = 2 \times 10^{21} \ \mathrm{m}^3$ and mobility $\mu \approx 0.63 \ \mathrm{m}^2/\mathrm{V \cdot s}$. The samples were glued to the surface of the epitaxial layer on silicon substrates with contacts for connection to the modulator circuit. The gluing thickness
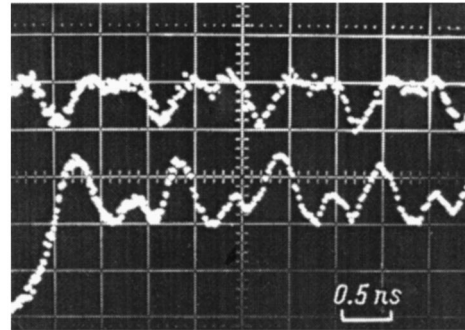
was $\sim 0.5 \ \mu$m, which provided heat removal and protected the GD surface from breakdown by the domain electric field. The GD contacts were made by annealing an Au−Ge eutectic deposited on the end faces of the crystals using a technology similar to Ref. 7. The typical dimensions of the GD crystal were $100 \times 80 \times 20 \ \mu$m (length×width×thickness); the threshold voltage of the GD was ~ 30 V and the threshold current ~ 100 mA. Figure 2 shows an oscillogram of the current oscillations for a 3G38 diode used for modulating the IL. The oscillogram was obtained in an electrodynamic test system with the diode biased by pulses from a transistor shaper with an $\sim 8 \ \Omega$ output resistance.

The modulation oscillograms in the current generator mode in the current-phase-matched GD–IL pair are presented in Figs. 3–5. A 3G38 Gunn diode ($R_0 = 240 \ \Omega$, threshold voltage 30 V) in the transit mode of 910 MHz oscillations generated pulses with $I_t = 96$ mA and $I_r = 52$ mA. The observed current-pulseduration was ~ 0.1 ns (Fig. 2). The N 29 injection laser ($\lambda = 0.85 \ \mu$m) had a threshold current $i_{th} = 48$ mA and a recommended working power of 5 mW at 67 mA. When the current was increased to 96 mA, the output power showed a tendency to saturate at 9 mW. The degree of modulation of the radiation and the load on the GD were changed by regulating the constant bias on the IL. For zero and positive (Fig. 3) biases the output radiation contains a constant component and the degree of modulation of the radiation is less than 1. The duration of the current pulses and the degree of modulation of the GD current are virtually independent of the bias on the IL. Thus the effect of the load is negligible in these cases. To achieve
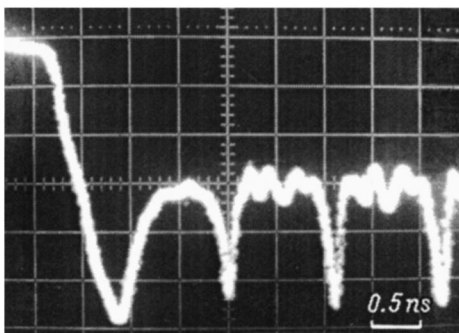

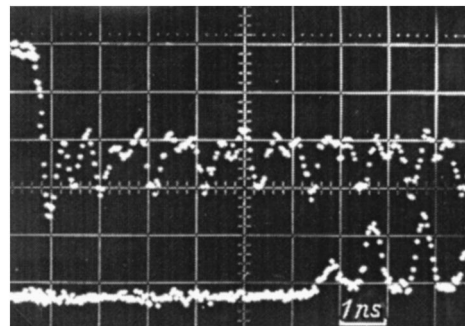FIG. 2. Test oscillogram of a modulating Gunn diode (17 mA/div along the ordinate, 0.5 ns/div sweep).

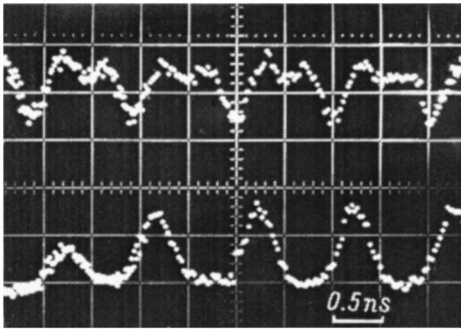FIG. 4. Same as in Fig. 3. Sweep 1 ns/div.

FIG. 5. Same as Fig. 3.

FIG. 6. Top trace — IL pump current (90 mA/div), bottom trace — photo-diode response (60 mV/div). Sweep 0.5 ns/div.

100% modulation a voltage in the blocking direction was fed to the bias circuit of the IL. This compensated the excess of the minimum current of the GD above the threshold current of the IL. The resistance of the IL and, correspondingly, the load for the GD increased in the process. This was reflected in the shape of the current pulses: Their duration and amplitude increased somewhat (Figs. 4 and 5).

The results of these experiments show that the current generator regime with modulation of the IL in the GD circuit was satisfactorily effectuated in the entire experimental range of biases on the IL. The bias was regulated from -2 (IL current 15 mA) up to the blocking voltage $+5$ V at the input of the bias circuit $R3$. There were no relaxation oscillations in the radiation from the IL. The form of the IL radiation corresponded to that of the modulation current. The second harmonic in the output radiation is sharper because the slope of the current–power characteristic of the IL is larger at low currents. The observed durations of the optical pulses are ~ 0.2 ns (rise time/fall time ~ 0.2 ns). We shall estimate the real pulse duration τ taking into account the rise time of the transient characteristic (time constant) of the photodetection system. For the estimate we employ Elmor's equation (the ''square root of the sum of squares rule'')[8] $\tau = \sqrt{\tau_0^2 - \tau_p^2 - \tau_s^2}$, where τ_0 is the observed duration, τ_p is the time constant of the photodiode, and τ_s is the time constant of the oscilloscope. According to the specification sheet of the S1-72 oscilloscope $\tau_s \simeq 0.1$ ns. For an LFD-2 photodiode a reasonable lower estimate is $\tau_p \geqslant 0.15$ ns. Then we obtain for the real pulse duration $\tau \leqslant 0.07$ ns.

The regime of modulation by voltage pulses is displayed in Figs. 6 and 7. In this regime a more powerful GD with threshold current 1.6 A and $\delta I = 0.7$ A was used. The required pump current was fed to the IL through a resistive divider. In the simplest case the IL was shunted by the load resistance R_L of the GD circuit, ranging from 1 to 4.5 Ω. The voltage generator regime obtained, since $R_L \ll R_l$, and the current through R_L was stabilized by the Gunn diode. It was determined that when relaxations are excited in the output radiation of the laser, oscillations are also observed in the pump current (Fig. 6). There is no relaxation in the voltage pulse on the GD load or on the IL-$R3$ circuit (Fig. 7). Since there are no reactive elements in the current dividing circuit, relaxation is caused by internal processes in the IL.[3] We note that some lasers had no relaxation ringing in any of the modulation regimes investigated.
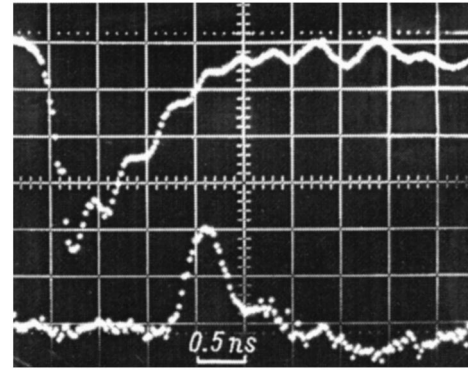
An important practical problem is obtaining stepped optical pulses. Such pulses are used, for example, for investigating the transient characteristics of photodetectors. Modulation by a current step was investigated in the current generator regime in a phase-matched GD–IL circuit. The current step was obtained by cutting off GD generation after several regular cycles at the moment of arrival of a domain at the anode. A $C1R1$ circuit connected in series with the GD was used for cutoff; as the capacitor charged up, the voltage on the GD decreased to a level somewhat below threshold. The maximum current hardly changed, since the current–voltage characteristic of the GD saturates in the subthreshold-voltage range. Modulation with a current-drop step δI was used to peak the leading edge of the optical pulse used to switch onthe IL. It was determined that the duration of the leading edge of the inherent optical switch-on pulse is itself determined mainly by the parameters of the IL structure. For the lasers investigated this was ~ 0.5 ns. For modulation by a current step the observed duration of the leading edge of the optical drop is ~ 0.2 ns (Fig. 8). The observed optical signal is actually a transient characteristic of the photodetecting system, since the duration of the leading edge of the pulse is essentially equal to the time constant of the photodetecting system (~ 0.18 ns). The true duration of the leading edge of the drop is estimated to be no greater than 0.07 ns.

The results of the present work confirm that the use of Gunn controlling components in integrated optoelectronic
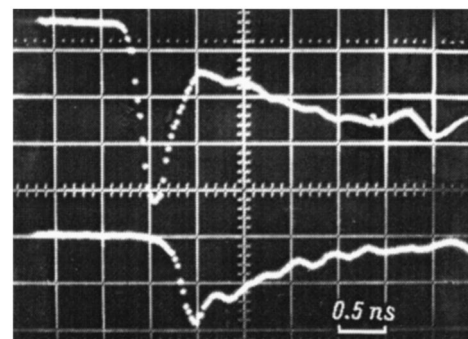

FIG. 7. Top trace — voltage on IL (1.2 V/div), bottom trace — IL current (180 mA/div).
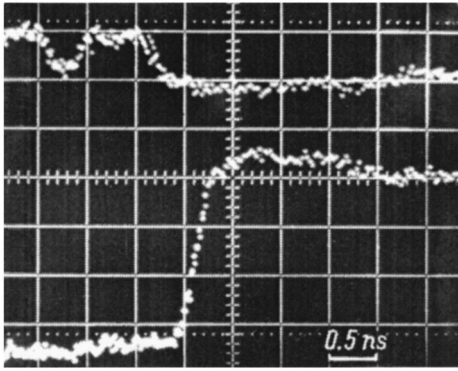
FIG. 8. Top trace — modulation current (23 mA/div), bottom trace — photodiode response (60 mV/div). Sweep 0.5 ns/div.

optical transmitters for measuring the characteristic (parasitic-free) response time of ILs. A phase-matched IL–GD pair can be used as a basis for building test optical-signal generators: stepped pulses (negative or positive step) and single pulses or series of short pulses with 10 ps rise/fall time. The pulse repetition frequency can reach tens of giga-hertz with high stability.

devices holds promises. Modulation of ILs by pulses from a GD current generator gives the maximum inherent fast-response of the IL. In practice it is also important that hard stabilization of the switching levels of the GD current protects ILs from being damaged by electrical noise in the power supply circuit.

The modulation regimes investigated can be used in fast

[1] M. E. Levinshteĭn, Yu. K. Pozhela, and M. S. Shur, *The Gunn Effect*, Sov. Radio, Moscow, 1975, 288 pp.
[2] M. S. Shur, *Modern Devices Based on Gallium Arsenide* [in Russian], Mir, Moscow (1991), 632 pp.
[3] W. Tsang (Ed.), *Semiconductor Injection Lasers*, Radio i Svyaz', Moscow (1990), 320 pp.
[4] H. W. Thim, L. R. Dawson, J. V. DiLorenzo *et al.*, IEEE Solid State Circuits Conf. Digest of Technical Papers (1973), pp. 92–93.
[5] H. Yanai, M. Yano, and T. Kamija, IEEE J. Quantum Electron. **QE-11**, 519 (1975).
[6] A. S. Logginov, K. Ya. Senatorov, V. E. Solov'ev *et al.*, Pis'ma Zh. Tekh. Fiz. **3**(24), 1326 (1977) [Sov. Tech. Phys. Lett. **3**, 547 (1977)].
[7] N. Yokoyama, S. Ohkawa, and H. Ishikawa, Jpn. J. Appl. Phys. **4**, 1071 (1975).
[8] E. I. Manaev, *The Principles of Radio Electronics*, Sov. Radio, Moscow, 1976, 480 pp.

Translated by M. E. Alferieff

Investigation of gallium arsenide single crystals with various crystallographic orientations, after implantation of silicon ions and pulsed photon annealing

S. V. Vasil'kovskiĭ, R. V. Konakova, and Yu. A. Tkhorik[†]

Institute of Semiconductor Physics, Ukrainian National Academy of Sciences, 252028 Kiev, Ukraine

M. P. Dukhnovskiĭ

Government Science and Industrial Association ''Istok,'' 141120 Fryazino, Moscow Region, Russia
(Submitted January 27, 1998)
Zh. Tekh. Fiz. **69**, 78–82 (May 1999)

The results of experimental investigations of gallium arsenide single crystals with the orientations (100), (311)A, (211)A, (111)A, and (221)A are presented. The crystals were doped with silicon ions on the Iolla-3M setup (ion energy 75 keV, ion beam density 1 μA/cm^2, implantation dose 1.2×10^3 cm^{-2}) at room temperature and annealed on the Impul's-5 setup at 950 °C. Raman scattering and low-temperature photoluminescence methods established that the highest electrical activity of the implanted silicon under identical implantation and annealing conditions obtains for (100) and (311)A gallium arsenide. In the process *n*-type layers are produced. © *1999 American Institute of Physics.* [S1063-7842(99)01305-7]

INTRODUCTION

Ion doping of semi-insulating GaAs with Si ions is now one of the principal technological methods for producing the working regions of *n*-type microwave transistors and fast integrated circuits.[1] Nonetheless, at present the physicochemical processes occurring here are not understood precisely. This is due to the complexity of the behavior of amphoteric Si impurity in gallium arsenide. Silicon can replace both Ga (*n*-type) and As (*p*-type), depending on the method used to grow GaAs and on the doping of GaAs,[2] the compensation of the substrate,[3] the presence of internal mechanical stresses in the substrate,[4] the annealing conditions,[1] and other factors. The effect of the GaAs substrate orientation on the characteristic features of silicon implantation in a GaAs lattice has been investigated in a number of works[5–7] using different deposition technologies. Photoluminescence (PL) methods have shown[2] that the homoepitaxial films obtained by molecular-beam epitaxy (MBE) on (211), (221), and (311) GaAs substrates were of much higher quality than those on the conventional (100) and (111) planes. We note that the literature contains virtually no information about the effect of the crystallographic orientation of the GaAs substrate on the characteristics of the Si-ion implanted layers. Therefore it was of interest to investigate the characteristic features of the formation of doped GaAs layers implanted with silicon ions in various crystallographic planes.

In the present work we used contact-free nondestructive optical methods of Raman scattering (RS) and low-temperature photoluminescence to study structural disordering processes, accompanying ion implantation and subsequent annealing, and the electrical activation of the dopant.

EXPERIMENTAL PROCEDURE

GaAs (AGChP-5) single crystals with (100), (311)A, (211)A, (111)A, and (221)A orientations were investigated. The crystals were cut from the same ingot. Before ion implantation the damaged layer was removed from the substrate surface by etching for 2 min in a 5:1:1 solution of H_2SO_4, H_2O, and H_2O_2 at temperature 60–70 °C. The implantation of silicon ions was performed on the Iolla-3M setup at room temperature with ion energy $E=75$ keV and ion beam density 1 μA/cm^2. The implantation dose was 1.2×10^{13} cm^{-2}. Annealing was performed in a pulsed heating regime, without protective coatings, on the Impul's-5 setup at 950 °C.[3] The temperature–time regime was as follows. A plate was held for 2 min at 200 °C, after which the temperature was raised to 950 °C in 0.5 min. Then the temperature was decreased to room temperature in 4 min.

The Raman scattering spectra were measured at $T=300$ K on an apparatus based on a DFS-52 spectrometer. The PL spectra were recorded at 4.2 K. The Raman scattering and PL signals were detected in the photon-counting mode using cooled FÉU-136 and FÉU-62 photomultipliers. Ar$^+$ and Kr$^+$ laser radiation ($\lambda=514.5$ and 647.1 nm, respectively) was used to excite the spectra. An ''in reflection'' experimental geometry was chosen with angle of incidence close to the Brewster angle. Polarization analysis of the scattered radiation was not performed. The spectral resolution was 1.5 cm^{-1} for the Raman scattering and 0.1 MeV for the PL spectra. The peaks of the RS bands were determined with an accuracy of 0.2 cm^{-1}.

RESULTS AND DISCUSSION

Raman scattering investigation of the structures. It is well known that two bands, due to scattering by longitudinal (*LO*) and transverse (*TO*) optical phonons, can appear in the first-order Raman scattering spectrum. According to the selection rules, depending on the crystallographic orientation, scattering by *LO* phonons ((111) plane), *TO* phonons ((110) plane), and *LO* and *TO* phonons ((111) plane) can
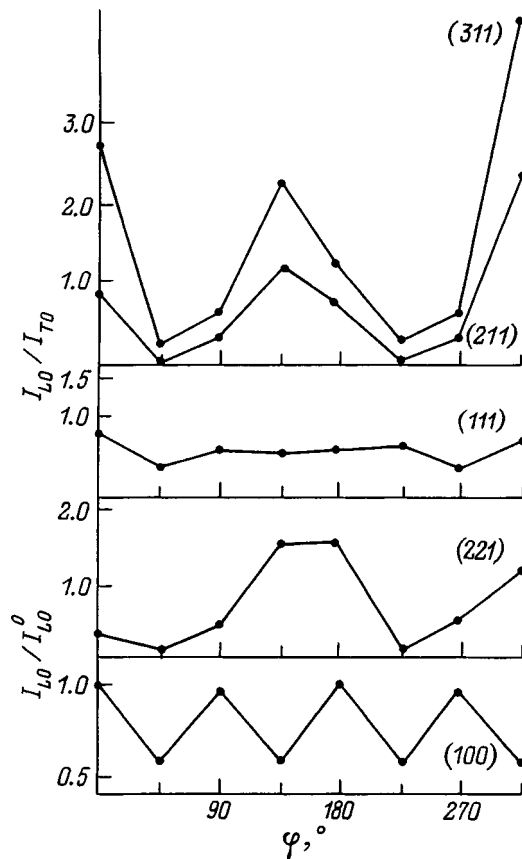
FIG. 1. Intensity ratios of the *LO* and *TO* phonon peaks versus the rotation angle φ of semi-insulating GaAs samples with different orientations around the normal to the given crystallographic plane.

FIG. 2. Half-widths Γ of the *LO* peaks (a) and *TO* peaks (b) of the initial (*1*) and implanted and then annealed (*2*) samples for the investigated orientations with excitation of RS by $\lambda = 514.5$ nm radiation; *3* — excitation by $\lambda = 647.1$ nm radiation.

occur in the experimental geometry employed.[8] The form of the Raman scattering spectrum for other planes is not clear a priori. It can only be expected that *LO* and *TO* phonon bands with different intensity ratios can appear. It is also not clear how implantation will change the Raman scattering spectrum in these cases. It should be noted that the GaAs surface manifests polarity with not only the (111) but also the (211), (311), (511), and so on orientations, the degree of polarity decreasing with increasing first index.[9]

The spectra of the initial samples contained intense symmetric peaks with half-widths $\simeq 3.5$ cm^{-1} at the frequencies 291.5 ± 0.2 and 267.5 ± 0.3 cm^{-1}. This corresponded to scattering by *LO* and *TO* phonons at the center of the Brillouin zone. The intensity ratios of the peaks in the spectra of the experimental samples were different. Figure 1 shows the intensity ratio of the *LO* and *TO* phonon peaks as a function of the rotation angle of the samples around the normal to a given crystallographic plane. The *TO* phonon peak was not present in the RS spectrum of the (100) sample. For this reason the intensity was normalized to the intensity of the *LO* peak with zero rotation angle of the sample. We note that rotation in this case is equivalent to rotation of the electric field vector **E** of the exciting radiation. The frequencies and half-widths of the peaks remained unchanged when the sample was rotated. As one can see from Fig. 1, the maximum redistribution of the intensities of the *LO* and *TO* peaks was observed in the spectra of the (311) samples, and
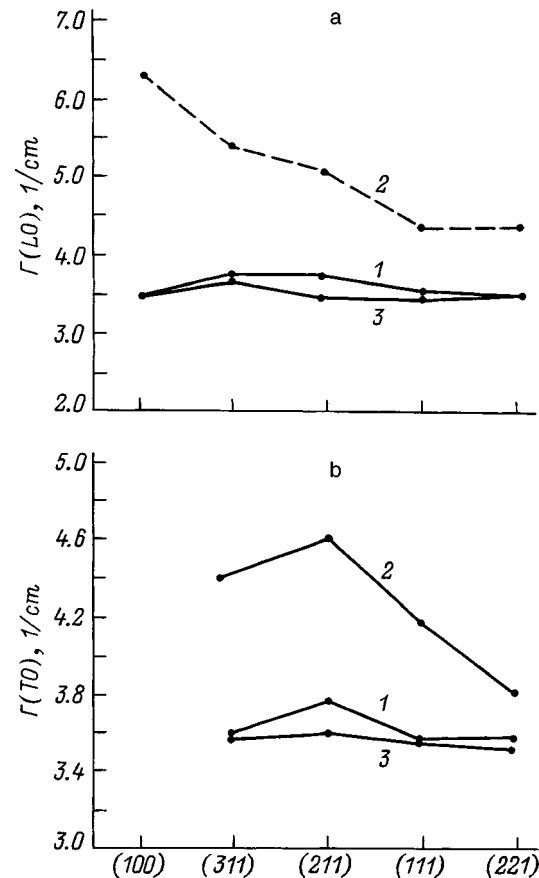
the minimum redistribution was observed for the (111) orientation. This can be explained by the fact that for scattering in the (111) plane the selection rules allow both peaks, while for the (100) plane only the *LO* phonon peak is allowed. The (311), (211), and (221) planes fall between these two cases.

In what follows we shall consider the effect of Si$^+$-ion implantation in GaAs with various crystallographic orientations on the RS spectra. When the RS spectra of implanted GaAs are excited with $\lambda = 647.1$ nm radiation, the broadening of the *LO* and *TO* peaks and their frequency position as compared with the initial GaAs remain virtually unchanged for all experimental samples. This shows that the implanted region in the semiconductor lies at a depth that is much smaller than the penetration depth of the probe radiation, and its contribution to the scattering cross section for a given wavelength is small.

When the Raman scattering is excited by radiation with a shorter wavelength $\lambda = 514.5$ nm, the relative intensity of the *LO* peak in the spectra is observed to decrease. This decrease is characteristic for all experimental samples. Moreover, the widths of the *LO* and *TO* peaks increase and the peaks shift in the high-frequency direction. The frequency shift and broadening were greater for the *LO* peaks. Thus the frequencies of the *LO* (*TO*) phonon peaks increased by 0.3 (0.2) cm^{-1} for (100), (311), and (211) orientations and by 0.6 (0.3) cm^{-1} for (111). Figure 2 shows the half-widths of

the LO and TO peaks of the initial and implanted samples for the orientations investigated. As one can see from Fig. 2, the half-width of the peak was maximum ($\simeq 6.5$ cm) for the (100) orientation and decreased with increasing polarity of the plane.

To analyze the RS spectra of implanted samples it is necessary to take into account the nonuniform depth distribution of the implanted ions. This distribution can be described in practice by two parameters: the average projected range R_p of the ions and the standard deviation ΔR_p of this range. Moreover, these parameters must be compared with the depth $d = 1/2\alpha$ of the layer probed by the radiation. This depth is determined by the wavelength of the exciting radiation because the value of the absorption coefficient $\alpha(\lambda)$ is different. For GaAs with Raman scattering excited by wavelengths $\lambda_i = 514.5$ nm and $\lambda = 647.1$ nm $d \simeq 100$ and 300 nm, respectively.[10] For implantation of Si^+ ions with energy $E = 75$ keV in GaAs $R_p = 61$ and $\Delta R_p = 36$ nm in the amorphous model.[11] Therefore in our case for $\lambda = 514.5$ nm the RS spectrum isdirectly determined by the implanted layer ($d \leq R_p + \Delta R_p$), and for $\lambda = 647.1$ nm it is also determined by the unimplanted layer of semi-insulating GaAs ($d > R_p + \Delta R_p$).

The observed change in the half-widths of the peaks for various orientations under excitation by 514 nm radiation could attest either to a different depth of the implanted region in the samples or a different electron density in the region probed by the laser radiation. This could be due to different degrees of electrical activation or self-compensation of the implanted impurity.

It is known that at least three factors can cause broadening of the peaks: an increase in the defect density,[12] a nonuniform distribution of the dopant and elastic stresses over the depth of the probed layer,[8] and the interaction of LO phonons with the plasma oscillations of the free carriers. The first two factors contribute to the broadening of the TO-phonon line. Apparently, two competing mechanisms operate in our case. The first one is due to disordering and the second one is due to the interaction of LO phonons with free carriers. The fact that after implantation the half-widths of the LO phonon peaks are greater than those of the TO phonon peaks confirms the influence of the second mechanism. Moreover, the increase in the half-width of the LO peak, the high-frequency asymmetry of this peak, and the decrease in its intensity could attest to an increase in the electron density in the samples. Therefore it can be inferred that the highest electrical activation of the implanted Si impurity under given implantation and annealing conditions obtains for the (100) and (311) planes.

Low-temperature PL investigation of the structures. We shall now consider the results obtained by PL at $T = 4.2$ K for (100), (311)A, (211)A, (111)A, and (221)A GaAs samples implanted with Si^+ ions and then annealed. The PL spectra of the initial and implanted samples are presented in Figs. 3 and 4, respectively. A complicatedband due to the photoluminescence of excitons bound on donors and acceptors appears in the spectra of all experimental samples in the region of the edge radiation (1.510–1.518 eV).[13] The most intense band is the $\simeq 1.49$ eV band. This band is due to radiative
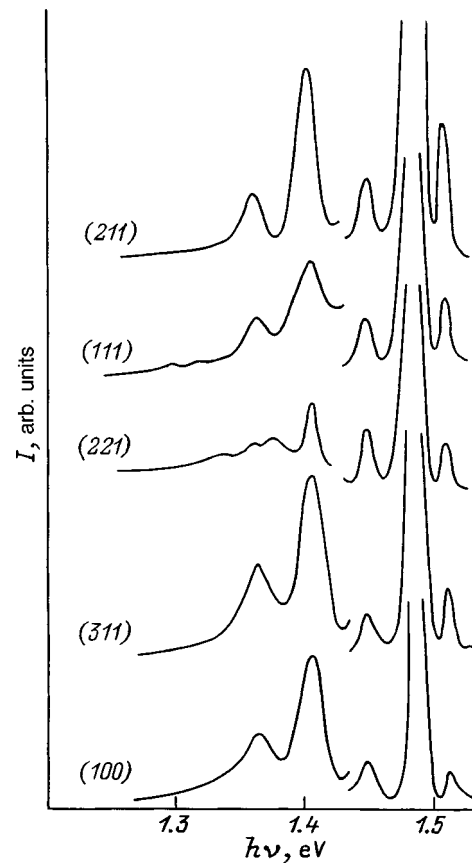

FIG. 3. PL spectra of semi-insulating GaAs samples with various orientations. The spectra were measured at $T = 4.2$ K.

recombination with the participation of carbon[14] and is characteristic for the PL spectra of GaAs, since carbon is one of the most common background impurities in gallium arsenide. A band with $h\nu \simeq 1.45$ eV also appears in the spectra of all samples. This band is probably due to radiative recombination corresponding to transitions including Si_{As}, (Si_{As}–V_{As}, Si_{As}–Ga_i).[15,16] A wide band due to radiative recombination with the participation of complex defect complexes, including V_{As} ($\simeq 1.405$ eV), is observed in the PL spectra.[17] A band with $h\nu \simeq 1.35$ eV was also observed in the PL spectra of the experimental samples. This band most likely is of a donor–acceptor nature and is due to an uncontrollable background impurity — copper, which is embedded in the gallium sublattice Cu_{Ga}.[18] The PL spectra of the implanted (100) and (311) samples contain a wide band with $h\nu \simeq 1.27$ eV. We shall consider it in greater detail. In the literature the $h\nu \simeq 1.27$ eV band is attributed to the formation of a large number of complicated defects which are due to a change in the charge composition of the atoms (Ga^{2-}/As^{2-} or Ga^0/As^0) or cluster defects.[19] However, in Ref. 20, where n-type Czochralski gallium arsenide doped with Si up to $\simeq 10^{18}$ cm^{-3} was investigated, the 1.27 eV band is attributed to radiative transitions with the participation of Si_{Ga}. The $h\nu \simeq 1.27$ eV band was not observed in the PL spectra of the unimplanted samples with any of the orientations that we studied.

We note that the spectra of the initial samples of semi-insulating GaAs (Fig. 3) with the orientations investigated
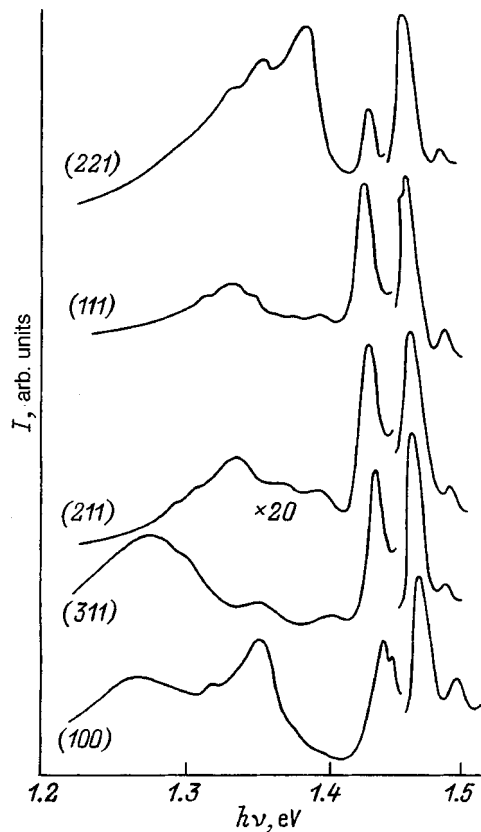
FIG. 4. Same as Fig. 3 but the samples were implanted with Si^+ ions followed by photon annealing.

show negligible differences in intensity and energy position (2–3) meV of the PL bands. We believe that this is due to the nonuniformity of the ingot and the polarization properties of the radiative centers.

Comparative analysis of the PL spectra of the initial and implanted samples showed that implantation followed by annealing of samples with all orientations studied does not greatly disrupt the structure of the GaAs crystal lattice. This is indicated by the presence of an excitonic band which is known to be very sensitive to the structural perfection of the crystal lattice, in all spectra. This correlates with the Raman scattering results, though it should be noted that the presence of bands with $h\nu \simeq 1.45$ and 1.49 eV in the spectra of the initial samples attests to the presence of background impurities in the initial semi-insulating GaAs. As noted above, these bands are due to radiative recombination with the participation of Si_{As} and carbon, respectively. The behavior of these bands in the spectra of samples with all orientations did not change much after implantation and annealing. We believe that this is because the implanted silicon predominantly occupies sites in the gallium sublattice or interstices, without influencing the Si_{As} concentration in the samples.

The differences between the PL spectra of implanted GaAs with different orientations, corresponding to the recombination radiation due to deep centers at energies 1.2–1.4 eV, turned out to be larger. As one can see from Fig. 4, the 1.40 eV band had the highest intensity in the spectra for (221) samples, while the 1.35 eV band had the highest intensity for (100). These bands are due to radiative recombina-

tion with the participation of arsenic vacancies. This confirms our conjecture that after implantation followed by annealing silicon occupies sites in the gallium sublattice. This is indicated more clearly by analysis of the PL spectra of (100) and (311) implanted samples, where a wide band at 1.27 eV, due to transitions with the participation of Si_{Ga}, appears after implantation. The absence of this band in the spectra of implanted and annealed samples with other orientations could attest to the fact that silicon is poorly incorporated into the gallium sublattice or the implanted atoms lie outside the layer producing the PL signal. We believe that the former is more likely, since no differences from the initial spectra are observed in the Raman scattering spectra of the implanted samples excited with $\lambda = 647.1$ nm radiation.

CONCLUSIONS

The experimental RS and PL data presented above show that the processes leading to defect formation accompanying implantation of Si in gallium arsenide with different crystallographic orientations followed by pulsed photon annealing differ substantially. This is evidently due to the difference in the mechanisms leading to the formation of the primary disordering tracks during the implantation process and to incorporation of the dopant by annealing. Moreover, it was shown that the degree of electrical activation of the impurity under identical implantation and annealing conditions is highest for the (100) and (311) orientations. The greatest broadening of the peak at the LO-phonon frequency and its high-frequency asymmetry were observed in the RS spectra of implanted and annealed samples for these orientations. A $h\nu \simeq 1.27$ eV band, due to radiative transitions with the participation of Si, was observed in the PL spectra in addition to an excitonic peak, indicating that the disruptions of the crystal structure of the samples are negligible.

In closing, we wish to thank G. G. Tarasov and G. A. Krysov for helpful remarks and a discussion of the results obtained in this work.

[1] J. S. Williams and J. M. Poate (Eds.), *Ion Implantation and Beam Processing* [Academic Press, New York (1984); Naukova Dumka, Kiev (1988), 369 pp.].

[2] P. N. Uppal, J. S. Ahearn, and J. W. Little, J. Vac. Sci. Technol. B **6**, 597 (1988).

[3] E. V. Gordienko, M. P. Dukhnovskiĭ, N. G. Lysenko *et al.*, Élektronnaya Tekhnika. Ser. 1, Élektronika SVCh, No. 6(390), 66 (1986).

[4] T. Otsuki, J. Appl. Phys. **61**, 928 (1987).

[5] T. P. Pearsall, R. E. Nachory, and J. R. Chelikowsky, Phys. Rev. Lett. **39**, 295 (1977).

[6] K. Kamon, M. Shimazu, K. Kimura *et al.*, J. Crystal Crow. **84**, 126 (1987).

[7] P. N. Uppal, J. S. Ahearn, and D. P. Muser, J. Vac. Sci. Technol. B **5**, 759 (1988).

[8] G. Burns, F. H. Dacol, and C. R. Wie, Solid State Commun. **62**, 449 (1988).

[9] W. I. Wang, E. E. Mendez, T. S. Kuan *et al.*, Appl. Phys. Lett. **47**, 826 (1985).

[10] M. Cargouri, B. Prevot, and C. Schab, J. Appl. Phys. **62**, 3902 (1987).

[11] A. F. Burlakov, F. F. Komarov, M. A. Kumakhov, and M. M. Temkin, *Tables of the Parameters of the Spatial Distribution of Ion-Implanted Impurities* [in Russian], Izd. BGU, Minsk, 1980, 374 pp.

[12] A. A. Maradudin, Solid State Physics **18**, 273 (1966); **19**, 1 (1966) [*Defects and Vibrational Spectrum of Crystals*, Mir, Moscow (1968), 432 pp.].

[13] J. Wagner and M. Ramsteiner, J. Appl. Phys. **62**, 2148 (1987).

[14] M. Maciaszek, D. W. Rogers, R. P. Bult *et al.*, Can. J. Phys. **67**, 384 (1989).

[15] A. V. Drazhan, Ukr. Fiz. Zh. **30**(3), 453 (1985).

[16] T. Hiramoto, Y. Mochizuki, and T. Ikoma, Jpn. J. Appl. Phys. **25**, L830 (1986).

[17] V. I. Gavrilenko, A. M. Grekhov, D. V. Korbutyak, and V. G. Litovchenko, *The Optical Properties of Semiconductors* [in Russian], Naukova Dumka, Kiev, 1987.

[18] É. É. Klotyn'sh, Izv. Akad. Nauk Latv. SSSR, Ser. Fiz. Tekh. Nauk, No. 4, 3 (1989).

[19] V. A. Bykovskaya, A. G. Bychkov, V. A. Zuev *et al.*, Poverkhnost', No. 10, 48 (1985).

[20] J. Pastrnak, J. Oswald, I. Gregora *et al.*, Phys. Solid State **11**, 345 (1989).

Translated by M. E. Alferieff

Measurement of the temporal and spectral characteristics of silicon *p–i–n* photodiodes in the soft x-ray range

A. V. Golubev, E. G. Pivinskiĭ, and V. V. Akulinichev

Institute of Analytical Instrumentation, Russian Academy of Sciences, 198103 St. Petersburg, Russia

A. A. Sorokin and S. V. Bobashev

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia

A procedure is briefly described for investigating the temporal and spectral characteristics of a soft x-ray detector in the range of photon energies from a few tenths of an electron-volt to more than a thousand electron-volts. The measured characteristics (signal rise time, time resolution, and absolute responsivity) are given, along with parameters (thicknesses of the contact layer, dead layer, and sensitive layer) determined from the measurement results for certain commercial brands of fast silicon *p–i–n* photodiodes from various manufacturers (Siemens, Hamamatsu, Motorola, and NIIIT/Moscow), which can be used in x-ray plasma diagnostic apparatus with a time resolution of 1 ns or better. © *1999 American Institute of Physics.* [S1063-7842(99)01405-1]

INTRODUCTION

Starting in the 1970s, silicon *p–i–n* photodiodes have played an active part in the time-resolved detection of radiation from a short-lived plasma in the soft x-ray range of the spectrum ($40\,\text{eV}<h\nu<6000\,\text{eV}$).[1] Many experimental papers have been published in the interim, reporting the utilization of photodiodes specially developed for the soft x-ray spectral range[2–4] as well as photodiodes designed for the visible range of the spectrum.[5,6] The widespread practical application of silicon *p–i–n* photodiodes has come about by virtue of several properties that set them apart in a favorable light from other semiconductor detectors. Among such properties are high time resolution, good responsivity in the soft x-ray spectral range for the photometry of pulsed plasma sources, a broad linear range, and relatively low cost.

To correctly interpret the results of pulsed plasma investigations, it is necessary to know the spectral and temporal characteristics of the photodiodes in the soft x-ray spectral range. The temporal characteristics dictate the capabilities of photodiodes in transmitting the temporal fine structure of radiation emitted by a pulsed source (e.g., Z-pinch and laser-produced plasma) in the form of output current pulse profiles. The spectral characteristics dictate the possibilities of using photodiodes for absolute soft x-ray photometry of a pulsed plasma.

Fast *p–i–n* photodiode diodes must have a sufficiently thin intrinsic conductivity *i* layer, small area of the active zone, and low resistance and capacitance of the *i* layer.[4] It is a well-known fact that the resistance and capacitance of the *i* layer can be significantly reduced by applying a bias voltage to the photodiode to generate an internal electric field exceeding $10^4\,\text{V/cm}$ in the structure. The thickness of the *i* layer, together with the area of the active zone, remains essentially the only factor that governs the response of a particular photodiode. The thickness of the *i* layer can be very accurately estimated from the experimentally measured risetime of the photodiode signal[7] produced in response to an input laser[7] or x-ray[3] δ pulse ($<0.2\,\text{ns}$).

The spectral characteristics of silicon *p–i–n* photodiodes in the soft x-ray (SXR) range depend mainly on the energy of electron–hole pair production in silicon, the transmission of soft x rays by the metal contact layer on the surface of the photodiode and the silicon ''dead layer,'' and the absorption of radiation in the sensitive *i* region of the semiconductor. A simple two-thickness model of the *p–i–n* photodiode[2,7–9] is used in conjunction with data on the absorption coefficients in the x-ray range of the spectrum (Ref. 11, Sec. 3) to calculate the responsivity in this case. The energy of electron–hole pair production energy in silicon has the same well-known value of $3.64\pm0.03\,\text{eV}$ (Ref. 9) for all silicon photodiodes. The thicknesses of the contact and dead layers can be determined from the experimentally measured absolute responsivity of the photodiode at individual spectral points of the soft x-ray range. Once the thicknesses of the absorbing layers are known, the spectral responsivity of the photodiode can then be calculated over the entire soft x-ray range of photon energies.

Generally speaking, the absolute spectrally selective calibration of any x-ray photodiode can be achieved by means of a primary standard detector.[10–14] Moreover, techniques are known for the absolute calibration of photodiodes using a primary radiation standard (e.g., synchronous radiation; Ref. 10) or by the self-calibration method using a synchronous radiation source.[15] The most reliable and accurate (within $<1\%$ error limits) spectrally selective calibration measurements are performed nowadays using a synchronous radiation source in conjunction with a cryogenic bolometer functioning as a primary standard.[12] In the calibration of fast

$p–i–n$ photodiodes, however, the use of a synchronous radiation source is hampered by the small area of such photodiodes in comparison with the width of the photon beam and by the unacceptably low signal-to-noise ratio at the output of the $p–i–n$ photodiode. A practical solution is to measure the absolute responsivityof $p–i–n$ photodiodes by means of a high-intensity pulsed (e.g., laser-plasma) soft x-ray source in combination with a primary standard detector.[13,16]

The objective of the present study is to investigate the temporal and spectral characteristics of $p–i–n$ photodiodes, which hold considerable promise from the standpoint of the diagnostics of short-lived, high-density plasmas in the soft x-ray range of the spectrum. The measurement procedure is described, and results are given from measurements of the temporal characteristics of four brands of photodiodes from different manufacturers: the SPPD11-04 from NIIIT (Scientific-Research Institute of Measuring Technology, Moscow), the Motorola MRD510, The Siemens BPX65, and the Hamamatsu S4753. We give a brief description of the procedure and summarize the results of systematic investigations ofthe spectral characteristics of the SPPD11-04 and MRD510 silicon photodiodes, which we have used for several years in experiments on various pulsed plasma objects.[16–21] We have also determined the spectral parameters of the BPX65 and S4753 photodiodes.

MEASUREMENTS OF THE TEMPORAL CHARACTERISTICS OF PHOTODIODES

We have determined the temporal characteristics of $p–i–n$ photodiodes from the measured responses of the photodiodes to soft x rays from a laser plasma generated by focusing short laser pulses onto a copper target contained in a vacuum chamber. The measurements were performed on a laser research facility built at the Institute of Analytical Instrumentation of the Russian Academy of Sciences, St. Petersburg.

The laser pulses were generated by the successive two-stage time compression of master oscillator pulses in stimulated Brillouin scattering[22] with subsequent amplification in two single-pass YAG:Nd^{3+} amplifiers. The master oscillator[23] comprised a passively Q-switched YAG:Nd^{3+} periodic-pulse laser ($\lambda = 1.064\,\mu$m, repetition rate 0.5 Hz, pulse duration 8 ns). The laser pulses generated at the system output had a duration of 0.12 ns and energy ~ 200 mJ, so that the radiation focused onto the target had a power density $\sim 10^{14}$ W/cm^2. The duration of the laser pulses, being the most important parameter for our measurements, was monitored by an Agat-SF-1 streak image tube, which has a time resolution of approximately 3 ps.

Commercial photodiodes from three manufacturers were chosen for the investigations: the Siemens BPX65, the Motorola MRD510, and the Hamamatsu S4753, which have been developed for the detection of visible radiation, along with one specimen of a specially constructed SPPD11-04 x-ray $p–i–n$ photodiode from NIIIT in Moscow.

The specification-sheet data on the area of the active zone and on the recommended supply voltage are known for all the diodes. The specifications for the responsivity and

time resolution in the visible range of the spectrum are also available for the BPX65, MRD510, and S4753 photodiodes (the S4753 has the best time resolution, approximately 0.7 ns, in the visible range). The gold contact layer on the surface of these photodiodes had a thickness of several tens of angstroms. We also know that the SPPD11-04 photodiode has an aluminum contact layer of thickness less than 0.1 μm on its working surface, the thickness of the dead layer is less than 0.3 μm, and the thickness of the sensitive zone is less than 80 μm.

Soft x-ray detectors were assembled from the photodiodes. This was done by placing each photodiode in a stainless steel casing and soldering its lead to a coaxial fifty-ohm output. The soft x-ray absorbing glass windows were first removed from the BPX65, MRD510, and S4753 photodiodes. The input windows of the detectors were coated with a fine nickel mesh with 67% transmission to suppress electromagnetic strays generated on the photodiodes by the formation of the laser plasma on the target. All the detectors were placed in the vacuum chamber at a distance of 10 cm from the target.

The input windows of the detector were covered with filters to separate out the required soft x-ray interval from the emission spectrum of the laser plasma while simultaneously preventing the pickup of stray visible and vacuum-ultraviolet (VUV) radiation from the plasma. The filters were thin metal films resting freely on the fine meshs. Three types of filters were used in our work to separate the radiation in three different spectral intervals: a copper film of thickness 0.46 μm, a silver film of thickness 0.27 μm, and an aluminum film of thickness 0.2 μm. All the filters were prepared and tested at the A. F. Ioffe Physicotechnical Institute (the transmission of the filters at a number of spectral points was measured on an RSM-500 spectrometer with an x-ray tube). The spectral transmission characteristics of the filters were calculated from test data on the thickness of the metal films and published data on the absorption coefficients.[11] The calculations showed that the copper filter could be used for the selection of radiation in the spectral interval from 500 eV to 933 eV with 10–35% transmission, the silver filter could be used to select radiation in the interval from 100 eV to 400 eV with 5–15% transmission, and the aluminum filter could be used to select radiation in the interval from 20 eV to 70 eV with 30–40% transmission. It is important to note that this partition into spectral intervals is to some degree conditional, since the transmission of all three filters increases rapidly and equally to 100% as the photon energy is increased in the range above 1000 eV. The photodiode signals therefore contained a high-energy component of the laser plasma radiation. The transmission of all three filters in the visible range of the spectrum was many orders of magnitude lower than in the x-ray range; at a wavelength of 650 nm, in particular, the measured transmission of all three filters was less than 10^{-6}%.

The electrical signals from the photodiodes were recorded by an SRG-7 wideband (5 GHz) storage oscilloscope. The time resolution $\tau_{0.5}$ of the oscilloscopic recording of the photodiode signals, defined as the full width at half maxi-
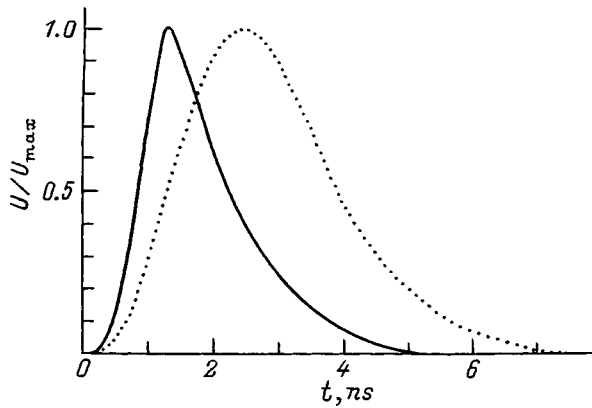
FIG. 1. Typical profiles of signals recorded by the SPPD11-04 photodiode using copper (solid curve) and aluminum (dotted curve) filters.

mum (FWHM) of the output current pulse in time units, can be determined from the equation[24]

$$\tau_{0.5} = (\tau_1^2 - \tau_2^2 - \tau_3^2 - \tau_4^2)^{1/2}, \qquad (1)$$

where τ_1 is the experimentally measured duration (FWHM) of the photodiode signal, τ_2 is the duration (FWHM) of the soft x rays incident on the photodiode, and τ_3 and τ_4 are the time resolutions of the coaxial transmission line and the cathode-ray tube (CRT) of the oscilloscope, respectively.

The duration of a soft x-ray pulse from a laser plasma source using a target with a high atomic number is known to be comparable with the duration of the heating laser pulse at photon energies of 1000 eV and higher.[25] We could assume on the basis of this fact that the soft x rays incident on the detector in measurements with a copper filter, which transmitted only radiation with photon energies above 500 eV, had a duration of approximately 0.12 ns. The time resolution of the experimental coaxial line (wave impedance 50 Ω, length 1 m) was 5×10^{-11} s, and the time resolution of the CRT was 5×10^{-11} s (Ref. 24). The experimentally measured durations of the photodiode signals (as an example, Fig. 1 shows typical signal profiles from a detector made from the SPPD11-04 photodiode with copper and aluminum filters on the input window) exceeded these values by more than an order of magnitude. We therefore assumed in measurements with a copper filter that $\tau_{0.5} = \tau_1$ within absolute error limits of 0.12 ns. In measurements using silver and especially aluminum filters the photodiode signals had much longer durations than in the measurements with a copper filter (Fig. 1). This difference is a consequence of the photodiodes detecting a lower-energy component of the laser

plasma radiation, with an emission time much longer than the duration of the plasma-generating laser pulse.

Parameters of major importance from the standpoint of photodiode applications in plasma diagnostics and photometry are the time resolution $\tau_{0.1}$, defined as the width (in time units) of the impulse response at 0.1 times its maximum, and the rise time τ_r of the photodiode signal from zero to the maximum. These parameters were also measured in experiments using a copper filter. The values of τ_r served as initial data for determining the thickness d_i of the sensitive zone of the $p-i-n$ photodiodes from the equation[7]

$$d_i = \tau_r \cdot V_{\mathrm{dh}}, \qquad (2)$$

where V_{dh} is the hole drift velocity in silicon.

The measured values of the time resolution $\tau_{0.5}$ and $\tau_{0.1}$, the signal rise time τ_r, and the thickness d_i of the sensitive zone for the four types of photodiodes are summarized in Table I.

MEASUREMENTS OF THE SPECTRAL RESPONSIVITY AND CALCULATIONS OF THE SPECTRAL CHARACTERISTICS OF PHOTODIODES

We know that the spectral responsivity $S(h\nu)$ of a silicon $p-i-n$ photodiode is traditionally defined as a quantity relating the photodiode current in amperes and the power in watts of monochromatic radiation incident on the photodiode. If the two-thickness model of the $p-i-n$ photodiode is adopted, its responsivity $S(h\nu)$ in units of A/W in the soft x-ray range of the spectrum can be written in the form[2,7,9,10]

$$S(h\nu) = S_{\max} \exp\{-\mu_c(h\nu)d_c\} \exp\{-\mu_{\mathrm{Si}}(h\nu)d_d\}$$
$$\times [1 - \exp\{-\mu_{\mathrm{Si}}(h\nu)d_i\}], \qquad (3)$$

where d_c, d_d and d_i are the thicknesses of the contact layer, the dead layer, and the sensitive layer, all of which absorb radiation, and μ_{Si} and μ_c are the linear x-ray absorption coefficients of silicon and the material of the contact layer, respectively; the maximum responsivity $S_{\max}$ is inversely proportional to the electron–hole pair production energy w in silicon ($w = 3.64$ eV; Ref. 8) and for all the silicon photodiodes has the same known value $S_{\max} = e/w = 0.275$ A/W, where e is the elementary charge.

It is important to note that this definition of the responsivity rests on the presumption that the electron–hole pair production energy in silicon does not depend on the energy in the soft x-ray range, and the distribution of the responsivity over the area of the active zone of the photodiode is known.[14]

TABLE I. Parameters of the investigated $p-i-n$ photodiodes.

Photodiode type	a, mm^2	Bias voltage, V	$\tau_{0.1}$, ns	$\tau_{0.5}$, ns	τ_r, ns	d_i, μm	d_c, μm	d_d, μm	$S^*_{\max}$, A·cm^2/W	$S_{\max}$, A/W
S4753	0.125	30	1.25	0.65	0.42	20	<0.01	<0.3	3.43×10^{-4}	0.275
MRD510	0.25	30	1.33	0.75	0.63	30	0.002	0.27	6.90×10^{-4}	0.276
BPX65	1.00	50	8.75	3.80	1.50	70	<0.01	>0.3	2.75×10^{-3}	0.275
SPPD11-04	5.00	160	3.30	1.40	1.00	50	0.025	0.20	1.38×10^{-2}	0.276
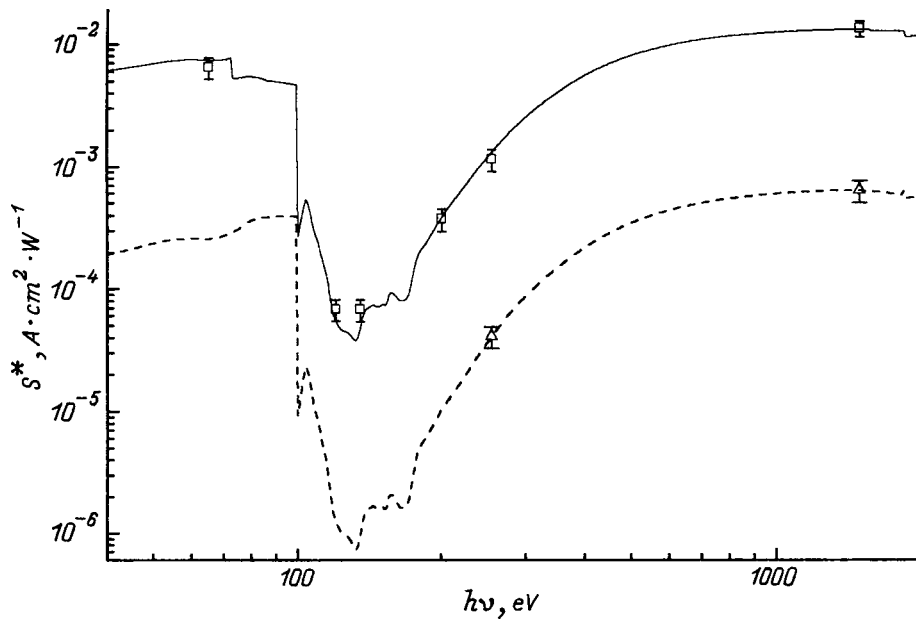
FIG. 2. Experimental values of the responsivity and spectral characteristics of the SPPD11-04 ($-\square-$) and MRD510 (- - -$\triangle$- - -) photodiodes.

In a number of practical situations it is useful to define the spectral responsivity as a quantity relating the photodiode current and the energy flux density of the radiation, expressed in W/cm², in the plane in which the photodiode is mounted. This situation often arises in calibration experiments and in the photometry of various pulsed plasma objects using $p-i-n$ photodiodes with small areas of the active zone, when the photodiode is completely inscribed within the calibration or investigated photon beam. In this case the spectral responsivity $S^*(h\nu)$ in units A·cm²/W can be expressed in the form

$$S^*(h\nu) = S^*_{max}\exp\{-\mu_c(h\nu)d_c\}\exp\{-\mu_{Si}(h\nu)d_d\}$$
$$\times [1 - \exp\{-\mu_{Si}(h\nu)d_i\}], \tag{4}$$

where S^*_{max} is a quantity that depends on the area a of the active zone of the photodiode; assuming a uniform distribution of the responsivity over the surface of the photodiode, we can write this quantity in the form

$$S^*_{max}[A \cdot cm^2/W] = S_{max} \cdot a = 0.275 \cdot a. \tag{5}$$

To determine the absolute responsivity of a photodiode at a spectral point of the soft x-ray range, it is necessary to know the thicknesses of the absorbing layers and the absorption coefficients of the layer materials. Moreover, in measurements of the energy flux density of soft x rays by fast $p-i-n$ photodiodes, which have a small area of the active zone and, generally speaking, an unknown responsivity distribution over the surface of the active zone, it is required to know the maximum responsivity S^*_{max}. The absorption coefficients are known for practically all materials.[11] The thicknesses of the absorbing layers and the values of S^*_{max} can be determined from Eq. (4) if the absolute responsivities of the photodiodes are known at at least a few spectral points.

In the present study measurements of the absolute responsivities S^*_{max} of the photodiodes have been carried out at several spectral points of the soft x-ray range at the Physicotechnical Institute using a bench and procedure described

earlier.[3,16,17] The bench included a laser plasma soft x-ray source, a photoionization quantometer functioning as a primary standard detector, and monochromatization channels utilizing flat and spherical multilayer x-ray mirrors (MLMs) and thin-film filters. Our procedure enabled us to measure the absolute responsivity of the photodiodes within 15–20% error limits, which depended mainly on the error of the absolute measurements of the soft x-ray fluxes by the photoionization quantometer.[13,16]

One distinctive feature of spectrally selective calibrations of silicon $p-i-n$ photodiodes in the photon energy range above 1000 eV is a strong dependence of the responsivity of these photodiodes on the photon energy. Owing to the absorption of radiation in the contact and dead layers, the responsivity of the photodiodes diminishes very rapidly as the photon energy decreases down to ~ 100 eV (the L absorption edge of Si). It is very difficult, therefore, to perform absolute spectrally selective calibration experiments in the spectral interval from 100 eV to approximately 300 eV. We have used spherical normal-incidence multilayer x-ray mirrors to obtain a high-intensity, quasi-monochromatic x-ray beam at the input to the photodiodes in the indicated spectral interval. By virtue of an increase in the solid angle of capture of laser plasma radiation we have succeeded in achieving a 20-fold increase in the radiation flux density on a calibrated $p-i-n$ photodiode, making it feasible to calibrate the MRD510 and SPPD11-04 photodiodes. However, even with the use of spherical multilayer x-ray mirrors the existing soft x-ray intensity was still inadequate for performing reliable spectrally selective calibrations of the S4753 and BPX65 photodiodes. To estimate the absolute responsivities of these photodiodes, we carried out comparative measurements of the S4753 and BPX65 photodiode output signals with the SPPD11-04 photodiodes, positioning all three detectors after a copper filter of thickness 0.46 μm.

The responsivities measured at several spectral points for the MRD510 and SPPD11-04 photodiodes are shown in

Fig. 2. Also shown in the figure are the spectral photodiode characteristics calculated from Eq. (4). The values of S^*_{max} and the thicknesses of the absorbing layers d_d and d_c, determined from the experimentally measured spectral responsivities, are shown in Table I. Also shown in the table are the values of S_{max} for the given detectors, calculated from Eq. (5), in which the values of a have been taken from specification-sheet data for the photodiodes. The tabulated thicknesses of the absorbing layers of the S4753 and BPX65 photodiodes have been estimated from our comparative measurements, and the values of S^*_{max} have been obtained from Eq. (5) using the known value of w and specifications for a.

DISCUSSION OF THE RESULTS

It has been observed experimentally that the maximum time resolution (see Table I) is attained at a supply voltage of $+30$ V for the MRD510 and S4753, at $+50$ V for the BPX65, and at $+160$ V for the SPPD11-04 photodiode. These values have been determined as the optimum for the given types of photodiodes. The duration of the photodiode signals remains essentially constant when the supply voltage is increased above the optimum, but the signals are observed to broaden considerably when the supply voltage is decreased. At the optimum supply voltages the electric field in the photodiodes increases to a value at which the carrier drift velocities in silicon attain maximum values $V_{de}=7.4 \times 10^6$ cm/s for electrons and $V_{dh}=4.8\times10^6$ cm/s for holes,[26] and the charges separate in minimum time. The measurements at the optimum supply voltages have been used as a basis for estimating the thickness d_i of the sensitive zone of the photodiodes, which is shown in Table I.

From the standpoint of photodiode applications in experiments to study subnanosecond temporal processes in a pulsed plasma, the greatest possibilities of all the investigated photodiodes are afforded by the S4753 and the MRD510, which have maximum time resolution at minimum supply voltage. However, their application as low-intensity radiation sources can be limited by their low responsivity in the subkilovolt spectral range. An example is found in Ref. 21, where an MRD510 photodiode was used for time-resolved measurements of soft x-ray fluxes from a high-density copper-aluminum plasma. In this work the measurements were successful only in the spectral range of photon energies above 350 eV.

The SPPD11-04 photodiode has the highest responsivity coupled with sufficiently high time resolution. The capabilities of this photodiode have been demonstrated in a series of experiments on the evolution of so-called hot spots in Z-pinches.[18–20] Measurements of line and recombination radiation in spectral intervals ranging from tens[20] to thousands[18,19] of electron-volts were measured in these experiments.

Compared with the other photodiodes investigated in this study, the BPX65 exhibits substantially inferior time resolution and low responsivity inthe subkilovolt region of the spectrum and therefore has only limited applications for short-lived plasma diagnostics.

[1] J. F. Cuderman and K. M. Gilbert, Rev. Sci. Instrum. **46**, 53 (1975).

[2] J. J. Hohlfelder, Adv. X-Ray Anal. **17**, 531 (1973).

[3] Z. A. Albikov, V. P. Belik, S. V. Bobashev *et al.*, in *Plasma Diagnostics* [in Russian], No. 6, edited by M. I. Pergament (Atomizdat, Moscow, 1989), pp. 48–52.

[4] O. Renner, E. Krousky, and L. Pina, Czech. J. Phys. **42**, 1 (1992).

[5] T. Jach and P. L. Cowan, Nucl. Instrum. Methods Phys. Res. **208**, 423 (1983).

[6] H. Fiedorowicz, P. Parys, and L. Ryc, Proc. Soc. Photo-Opt. Instrum. Eng. **1140**, 91 (1989).

[7] L. Pina, Czech. J. Phys., Sect. A **35**, 363 (1985).

[8] D. M Corallo, D. M. Creek, and G. M. Murray, J. Phys. E **13**, 623 (1980).

[9] F. Scholze, H. Rabus, and G. Ulm, Appl. Phys. Lett. **69**, 2974 (1996).

[10] E. Tegeler, Phys. Scr. **T31**, 215 (1990).

[11] B. L. Henke, E. M. Gullikson, and J. C. Davis, At. Data Nucl. Data Tables **54**, 181 (1993).

[12] H. Rabus, V. Persch, and G. Ulm, Appl. Opt. **36**, 5421 (1997).

[13] S. V. Bobashev and L. A. Shmaenok, Rev. Sci. Instrum. **51**, 16 (1981).

[14] E. M. Gullikson, R. Korde, L. R. Canfield *et al.*, J. Electron Spectrosc. Relat. Phenom. **80**, 313 (1996).

[15] M. Krumrey and E. Tegeler, Rev. Sci. Instrum. **63**, 797 (1992).

[16] S. V. Bobashev, G. S. Volkov, A. V. Golubev *et al.*, Phys. Scr. **43**, 356 (1991).

[17] S. V. Bobashev, A. V. Golubev, D. A. Mosesyan *et al.*, Zh. Tekh. Fiz. **65**(10), 62 (1995) [Tech. Phys. **40**, 1011 (1995)].

[18] M. Hebach, D. Simanovskii, S. Bobashev *et al.*, Plasma Sources Sci. Technol. **2**, 296 (1993).

[19] A. Krejci, J. Raus, V. Piffl *et al.*, IEEE Trans. Plasma Sci. **PS-21**, 584 (1993).

[20] A. V. Golubev and A. Krejci, *Soft X-Rays of Nitrogen Z-Pinch*, Research Report IPPCZ-352 (Prague, 1995), 18 pp.

[21] A. A. Sorokin, L. A. Shmaenok, V. P. Velik *et al.*, Pis'ma Zh. Tekh. Fiz. **22**(4), 15 (1996) [Tech. Phys. Lett. **22**, 272 (1996)].

[22] E. Pivinskii, V. Akulinichev and V. Gorbunov, Proc. Soc. Photo-Opt. Instrum. Eng. **2986**, 231 (1997).

[23] V. V. Akulinichev, M. E. Mavrichev, and E. G. Pivinskiĭ, Opt. Spektrosk. **76**, 681 (1994) [Opt. Spectrosc. **76**, 611 (1994)].

[24] L. A. Novitskiĭ and B. M. Stepanov, in *Photometry of Fast Processes* [in Russian], Mashinostroenie, Moscow (1983), 296 pp.

[25] G. L. Stradling, D. T. Attwood, and R. L. Kauffman, IEEE J. Quantum Electron. **QE-19**, 604 (1983).

[26] Y. K. Akimov (Ed.), *Semiconductor Detectors in Experimental Physics* [in Russian], Énergoatomizdat, Moscow (1989), 344 pp.

Translated by James S. Wood

Ferroelectric materials for dynamic-memory integrated circuits

B. M. Gol'tsman and V. K. Yarmarkin

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia

The possibilities of using ferroelectric materials for new generations of integrated circuits for high-density dynamic memory (up to 1 Gbit per crystal) are discussed. The correspondence of the specific capacitance and leakage currents of thin film ferroelectric capacitors to the requirements for integrated circuits with various information capacities is examined. It is shown that the capacitance–voltage characteristic of the ferroelectric strongly influences the specific capacitance and the rate of decrease of the voltage across the capacitors when they are discharged in the process of storing information. The prospects for increasing the specific capacitance of memory capacitors using relaxor ferroelectrics are examined. © *1999 American Institute of Physics.* [S1063-7842(99)01505-6]

The advancement of computer technology in the last ten years has been driven by the increase in the information capacity of integrated circuits used in the dynamic memory of working storage. Increasing information capacity involves a many-fold increase in the number of memory capacitors in an integrated circuit, which is accomplished by decreasing the area of the capacitors to less than 1 μm^2, increasing the area of the integrated circuit, and decreasing the gaps between the capacitors. Since the capacitance of the capacitors should not decrease much in so doing, it is necessary to increase sharply their specific capacitance C_s. According to estimates,[1] for integrated circuits with an information capacity of 64, 256 Mbit, and 1 Gbit the values of C_s should be, respectively, 23, 35, and 100 fF/μm^2 (for the simplest and cheapest planar construction of the capacitors). As the generations of integrated circuits in dynamic memory are replaced, the working voltage U decreases as a result of improvements in the technology for fabricating memory-cell transistors: 64 Mbit — 3.3 V; 256 Mbit — 2.5 V; 1 Gbit — 1.6 V. The leakage current density j of the capacitors should not exceed 10^{-7} A/cm^2 in all cases.

It is desirable to use thin-film ferroelectrics, whose dielectric constant ε can reach several thousand, to increase the specific capacitance of the memory capacitors. To obtain the required specific capacitances it is necessary to use ferroelectric films with thickness d not exceeding 100–200 nm. To increase the stability of the capacitors and eliminate the fatigue of the ferroelectric arising under repeated switching of the polarization, due to the motion of domain walls, materials in the paraphase at the working temperature or in the ferroelectric phase with narrow hysteresis loops are preferable.[1] Examples of such materials, which recently have been under intensive study for use in memory capacitors, are the paraelectrics SrTiO$_3$ (STO) and (Ba$_x$Sr$_{1-x}$)TiO$_3$ (BST) with $x \leq 0.7$ and $T \geq 0$,[2,3] and ferroelectrics with narrow hysteresis loops (Pb$_{1-x}$La$_x$)(Zr$_{1-y}$Ti$_y$)O$_3$ (PLZT) with $x = 0.09$ and $y = 0.35$ and Pb(Zr$_{1-x}$Tr$_x$)O$_3$ (PZT) with $x \simeq 0.5$ and film thickness less than 100 nm (Ref. 4) (the narrowing of the hysteresis loop in such thin films is due to the influ-

ence of size effects associated with the fine-grain structure of the films, intermediate layers at the film–electrode interfaces, mechanical stresses, and so on on their dielectric properties).

In the present paper we examine for a variety of ferroelectrics the correspondence of the parameters C_s and j of the film capacitors noted above to the requirements. We take into consideration the fact that, since large values of C_s are achieved with thin films, strong electric fields (up to 1 MV/cm) operate in the films at working voltages of 1.6–3.3 V. The effect of these films on C_s and on the rate of decrease of the voltage on a capacitor as it discharges in the process of storing information is examined. The prospects for increasing C_s of capacitors with various ferroelectrics by decreasing the thickness and increasing the permittivity of the films are assessed.

Figure 1a shows the capacitance C_s versus the voltage on a capacitor and versus the thickness of STO and BST films. It is evident that as U increases, C_s decreases approximately linearly. This is due to the characteristic, for these materials, decrease of ε with increasing electric field intensity.[5–7] It is also evident from Fig. 1a (for STO) that C_s increases more slowly with decreasing d than follows from the expression for the capacitance of a flat capacitor ($C_s \propto 1/d$). This could be due to an increase in the volume fraction of defective layers at the film–electrode boundaries as well as a decrease in the permittivity of the film as a result of size reduction of the crystal grains and an increase of the electric field intensity.

According to the data in Fig. 1a, 92 nm thick STO films approximately meet the requirements for C_s for 64 Mbit information capacity ($C_s = 20$ fF/μm^2 at $U = 3.3$ V). The leakage current density of the capacitor is $j < 10^{-8}$ A/cm^2 (Ref. 8). It is also evident that at a voltage of 3.3 V the field acting in the film substantially decreases C_s (by 30%). The main reserve for increasing C_s of STO-based capacitors is to decrease d within admissible values $j < 10^{-7}$ A/cm^2, since there are few possibilities for increasing ε of STO films by improving their fabrication technology: For relatively thick STO films ε reaches 300,[9] which is close to the values for
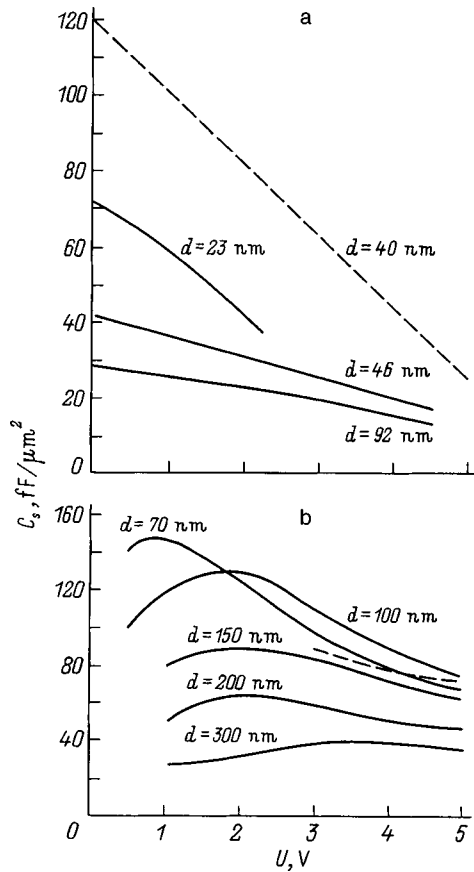
FIG. 1. Specific capacitance versus voltage for memory capacitors based on SrTiO$_3$ (Ref. 5) (solid line), (Ba$_{0.5}$Sr$_{0.5}$)TiO$_3$ (Ref. 11) (dashed line) (a) and Pb(Zr$_{0.5}$Ti$_{0.5}$)O$_3$ (Ref. 4) (solid line) and (Pb$_{0.91}$La$_{0.09}$)(Zr$_{0.65}$Ti$_{0.35}$)O$_3$ (Ref. 4) (dashed line) (b).

bulk samples. Likewise, there are no substantial reserves for increasing ε of thin films by improving the technology, since even now ε decreases by only 40% when the film thickness decreases from 150 to 20 nm (Fig. 2).[10–16]

The value $C_s = 120$ fF/μm^2 with $U \simeq 0$ was obtained in BST films with $x = 0.5$ and $d = 40$ nm, and the sensitivity of C_s to the bias voltage was estimated.[11] Using this estimate and assuming a linear dependence $C(U)$ (Fig. 1), C_s can be determined for values of U corresponding to various information capacities of memory integrated circuits. An estimate
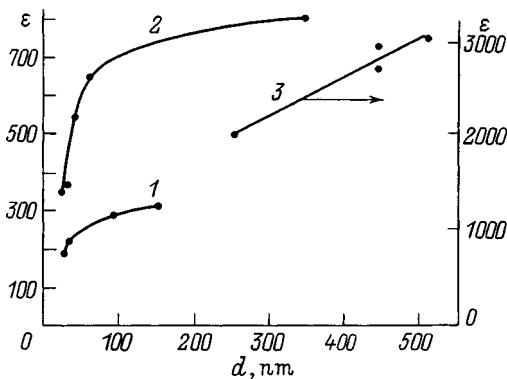

FIG. 2. Dielectric constant of SrTiO$_3$ (*1*), (Ba$_{0.5}$Sr$_{0.5}$)TiO$_3$ (*2*), and 0.9 Pb(Mg$_{1/3}$Nb$_{2/3}$)O$_3$–0.1 PbTiO$_3$ (*3*) films versus thickness.[5,10–16]

gives $C_s = 71$ fF/μm^2 at $U = 2.5$ V and $C_s = 90$ fF/μm^2 at $U = 1.6$ V. Since $j < 10^{-8}$ A/cm^2 and $U = 1.6$–2.5 V,[11] BST films meet the requirements for C_s, U, and j_L for 256 Mbit and are close to (with respect to the value of C_s) the requirements for 1 Gbit. As one can see from the data presented, a field acting in the film decreases C_s by 40% at $U = 2.5$V and by 25% at $U = 1.6$ V. Here, just as for STO films, C_s can be increased by decreasing the film thickness within admissible values of the leakage current density.

In contrast to STO, for BST the transition from bulk to film samples, just as decreasing d from 500–1000 to 20 nm, sharply decreases ε (by at least a factor of 3). It is obvious that the specific nature of the film formation process, consisting of simultaneous formation and growth of a large number of crystalline grains, gives rise in the case of BST to greater (compared with STO) degradation of the structural perfection in both the interior volume of the films and in the regions near the electrodes. Degradation of the structural perfection of films is taken to mean the development of intercrystallite boundaries in connection with the size reduction of the grains and blocks, an increase in the density of impurities and vacancies, other point defects, and dislocations, and an increase in the mechanical macro- and microstresses. The strong structural degradation of BST films can be explained by the more complicated composition of the solid solution BST as compared with the individual chemical compound SrTiO$_3$ (STO). This suggests that in the case of BST there are reserves for increasing C_s by improving the film fabrication technology, leading to higher values of ε.

It should be noted that as the structural perfection of STO and BST films improves, the dependence $C(U)$ becomes sharper. Thus, at $U = 0.8$ V ($E = 210$ kV/cm) the value of C_s in 38 nm thick BST films decreases by 15%,[17] whereas for 1000–1200 nm thick films with the same composition and a coarse-grain structure (annealing at 1150 °C) C_s decreases by 30% (for the same value of E).[7] This decreases the increase in C_s of BST films due to the improvement of their structural perfection, and it increases the rate of decrease of the voltage on a capacitor as a result of discharging of the capacitor during storage of information. Indeed, a decrease in voltage during the discharging of the capacitor increases the capacitance, which in turn additionally decreases the voltage, the capacitance once again increases, and so on.

Since the rate of voltage drop on the capacitor should not be very high, we shall determine the change in time t of the initial voltage U_0. To simplify the calculation, we set $j = U/R$, where $R =$ const. Following Fig. 1a we shall consider the linear dependence

$$C(U) = C_0 - BU, \tag{1}$$

where B is a constant coefficient.

Equating the derivative dQ/dt ($Q = CU$ is the charge on the capacitor) to the leakage current, we obtain

$$dQ/dt = C dU/dt + U dC/dU \, dU/dt = -U/R. \tag{2}$$

From Eqs. (1) and (2) we have

$$dt = 2BR \, dU - C_0 R \, dU/U. \tag{3}$$

Integrating Eq. (3), we obtain

$$U/U_0 \exp[2B(U_0 - U)/C_0] = \exp(-t/RC_0). \quad (4)$$

The relation (4) makes it possible to determine the extent to which the discharge time will decrease under the influence of the capacitance $C(U)$ to a given value of U or how much the voltage drop U will increase over a given discharge time. An estimate made for $C(U)$ for BST films (Fig. 1a) showed that the time in which U of a capacitor charged to $U_0 = 2.5$ V decreases by 10% is decreased by a factor of 4 as a result of the dependence $C(U)$, and for a fixed discharge time the decrease reaches 30% instead of 10% if $C(U)$ is neglected.

The character of the dependence $j(U)$ also affects the rate of decrease of the voltage during the discharging of a capacitor. This dependence is determined by the structure and composition of the film as well as by the potential barriers at the film–electrode boundaries. As a rule, the dependence $j(U)$ deviates from Ohm's law in the direction of a larger decrease of the leakage current with decreasing voltage. This slows down the discharging of the capacitor. To estimate the effect of the dependence $j(U)$ on the change in the voltage on the capacitor during discharging we shall compare the average values of the leakage currents over a time of a 10% decrease in voltage (from $U = 2.5$ V) for an Ohmic dependence $j(U)$ and for the stronger dependence presented in Ref. 11. In the latter case the average leakage current and therefore the rate of discharging of the capacitor are 10% less then for an Ohmic dependence, i.e., the voltage decreases by 9% during discharging. Therefore the effect of a deviation of $j(U)$ from Ohm's law is negligible.

Figure 1b shows C_s versus U for PZT ($x = 0.5$) and PLZT ($x = 0.09$, $y = 0.35$) films. The curves were constructed based on the data of Ref. 4. In contrast to the corresponding data for STO and BST, these curves for PZT possess a maximum, which probably reflects the fact that the films are in the ferroelectric phase and the fields acting in the films are close to the coercive values. The distinction from STO films is that, as one can see from Fig. 1, in the region of practical interest $U = 1.6$–2.5 V, C_s for 100–300 nm thick PZT films changes as $C_s \propto 1/d$. This is explained by the weak dependence $\varepsilon(E)$ in the indicated voltage range (the intensification of the field with decreasing d does not affect the capacitance much). For large U the field-induced decrease of the capacitance intensifies, which decreases the dependence $C_s(d)$. As a result, the curve $C_s(U)$ and $d = 70$ nm for $U > 1.7$ V lies below the curve for $d = 100$ nm. On account of the weak voltage dependence $C(U)$ for working voltages of 1.6–2.5 V, in the case of PZT the above-noted increase of the rate of voltage decrease during the discharging of a capacitor should be negligible because of the growth of C_s with decreasing U.

For PZT films with $U = 1.6$ V (according to the requirements for 1 Gbit information capacity) and $d = 70 - 100$ nm $C_s = 130$ fF/μm^2,[4] which is greater than the minimum value 100 fF/μm^2 required for this information capacity. However, j exceeds 10^{-7} A/cm^2 ($j = 5 \times 10^{-7}$ A/cm^2) in this case. There is a possibility of decreasing j by increasing d and thereby approaching the 1 Gbit requirements.

It can be assumed that the parameters of PLZT-based capacitors are also close to the requirements for 1 Gbit in-

formation capacity. For $U = 1.6$ V $j \simeq 10^{-7}$ A/cm^2 (for 150 nm thick films[4]), and extrapolating the curve $C_s(U)$ (Fig. 1b) gives $C_s \simeq 100$ fF/μm^2 at this voltage.

It should be noted that the values of the parameters C_s, U, and j do not cover all requirements for memory capacitors, and there can be problems in meeting other requirements. For example, in Ref. 18 ε of BST films was observed to decrease by 10% after 10^9 charging–discharging cycles with Ir electrodes and by 30% after 10^8 cycles with Pt electrodes, which is much less than the required 10^{12} cycles.[19] Therefore it is necessary to consider the possibility of using other ferroelectric materials besides those considered above.

Compounds and their solutions from an extensive family of ferroelectrics with perovskite structure and a diffuse phase transition (so-called relaxor-type ferroelectrics) are promising. These materials were discovered and investigated in detail by G. A. Smolenskiĭ and his colleagues.[20] These materials are distinguished by high values of ε and a narrow dielectric hysteresis loop.

The little published information that is available on the properties of thin films consisting of the materials indicated above concerns comparatively thick films ($d \geqslant 250$ nm) so that even though the values of ε in these films are quite high the values of C_s for capacitors based on them are much lower than for capacitors based on BST, PZT, and PLZT. Thus for Pb(Mg$_{1/3}$Nb$_{2/3}$)O$_3$ (PMN)-based capacitors $C_s = 67$ fF/μm^2 for $d = 500$ nm and $\varepsilon = 3800$,[21] for the solid solution 0.9 PMN—0.1 PbTiO$_3$ (PT) $C_s = 70$ fF/μm^2 for $d = 250$ nm and $\varepsilon = 2000$,[16] and for Pb(Sc$_{1/2}$Ta$_{1/2}$)O$_3$ (PST) $C_s = 15$ fF/μm^2 for $d = 2700$ nm and $\varepsilon = 4500$.[22] These values show that ε in these materials decreases substantially on switching from bulk samples to thin films (for example, in bulk samples of the solid solution 0.9 PMN–0.1 PT $\varepsilon = 30000$[23]), which indicates the existence of large reserves for increasing C_s of capacitors by increasing ε of the films by improving their production technology. The value C_s can also be increased by decreasing d.

CONCLUSIONS

1. An estimate of the specific capacitance C_s of memory capacitors based on ferroelectrics, taking account of their capacitance–voltage characteristic and leakage currents, shows that the values of C_s for capacitors with STO and BST films meet the requirements for integrated circuits for 64 and 256 Mbit dynamic memories, respectively. The values of C_s for capacitors with PZT and PLZT films are close to the requirements for 1 Gbit information capacity.

2. Capacitors with films of relaxor ferroelectrics (PMN, PMN–PT, PST) are at present inferior with respect to C_s to BST-, PZT-, and PLZT-based capacitors, but they have substantial reserves for increasing their specific capacitance both by decreasing the film thickness and by decreasing the large gap between the values of ε for films and bulk materials. Such reserves are also available for BST-based capacitors.

3. Calculations show that in capacitors with films in the paraphase (for example, BST), the rate of voltage decrease during discharging increases substantially because of the

strong capacitance–voltage dependence. In capacitors with films in the ferroelectric phase (PZT) this effect should be weak for fields close to the coercive fields.

[1] D. E. Kotecki, Integrated Ferroelectrics **16**, 1 (1997).
[2] B. Jaffe, W. Cook, and H. Jaffe, *Piezoelectric Ceramics* [Academic Press, New York (1971); Mir, Moscow (1974), 288 pp.].
[3] V. V. Lemanov, E. P. Smirnova, P. P. Syrnikov, and E. A. Tarakanov, Phys. Rev. B **54**, 3151 (1996).
[4] R. E. Jones, P. D. Maniar, A. C. Campbell *et al.*, Integrated Ferroelectrics **5**, 235 (1994).
[5] K. Abe and Sh. Komatsu, Jpn. J. Appl. Phys. **32**, L1157 (1993).
[6] B. M. Gol'tsman, A. I. Dedyk, V. V. Lemanov *et al.*, Fiz. Tverd. Tela (St. Petersburg) **38**, 2493 (1996) [Phys. Solid State **38**, 1368 (1996)].
[7] B. M. Gol'tsman, V. V. Lemanov, A. I. Dedyk *et al.*, Pis'ma Zh. Tekh. Fiz. **23**(15), 46 (1997) [Tech. Phys. Lett. **23**(8), 594 (1997)].
[8] K. Abe and Sh. Komatsu, Jpn. J. Appl. Phys. **31**, 2985 (1992).
[9] S. Liang, C. S. Chem, Z. Q. Shi *et al.*, Appl. Phys. Lett. **64**, 3563 (1994).
[10] W.-J. Lee, H.-G. Kim, and S.-G. Yoon, J. Appl. Phys. **80**, 5891 (1996).
[11] B. A. Baumert, L.-H. Chang, A. T. Matsuda *et al.*, Integrated Ferroelectrics **17**, 165 (1997).
[12] C. S. Hwang, S. O. Park, H.-J. Cho *et al.*, Appl. Phys. Lett. **67**, 2819 (1995).
[13] C. S. Hwang, C. S. Kang, H.-Ju. Cho *et al.*, Integrated Ferroelectrics **12**, 199 (1996).
[14] Sh. Nagakari, K. Kamigaki, and Sh. Nambu, Jpn. J. Appl. Phys. **35**, 4933 (1996).
[15] K. R. Udayakumar, J. Chen, V. Kumar *et al.*, in *Proceedings of the 7th International Symposium on Applications of Ferroelectrics*, Urbana–Champaign, 1990, pp. 744–746.
[16] L. F. Francis and D. A. Payne, *ibid.,* pp. 263–266.
[17] H.-J. Cho, C. S. Kang, C. S. Hwang *et al.*, Jpn. J. Appl. Phys. **36**, L874 (1997).
[18] S. Y. Cha, B.-T. Jang, D.-H. Kwak *et al.*, Integrated Ferroelectrics **17**, 187 (1996).
[19] D.-S. Paik, H.-Y. Shin, H.-W. Choi *et al.*, Ferroelectrics **200**, N 1 (1997).
[20] G. A. Smolenskiĭ, V. A. Bokov, V. A. Isupov *et al.*, *Physics of Ferroelectric Phenomena* [in Russian], Nauka, Leningrad, 1985.
[21] C. Tantigate, J. Lee, and A. Safari, Appl. Phys. Lett. **66**, 1611 (1995).
[22] A. Patel, N. Shorrocks, and R. Whatmore, Ferroelectrics **134**, 343 (1992).
[23] E. P. Smirnova, O. V. Rubinshtein, and V. A. Isupov, Ferroelectrics **143**, 263 (1993).

Translated by M. E. Alferieff

On the mechanism of the photochromic process in copper-halide photochromic glasses

V. A. Reznikov, L. E. Solov'ev, and V. E. Kholmogorov

St. Petersburg State University, 199034 St. Petersburg, Russia

It is demonstrated that copper oxides (specifically cuprous oxide) play an important role in the photochromism of standard FKhS-7 glasses. The absorption spectra of the FKhS glasses are compared with the spectrum of glasses with microscopic cuprous oxide particles and the spectra of cuprous oxide deposited on various substrates. Solid-phase photochemical reactions in heterosystems are shown to be responsible for photochromism. © *1999 American Institute of Physics.* [S1063-7842(99)01605-0]

As shown in Refs. 1 and 2, at moderate and low irradiation doses of $\lambda \geq 300$ nm light the induced absorption spectrum ΔD in FKhS-7 glasses and the absorption spectrum of nanosize Cu_n copper particles in a dielectric matrix correlate. The observed displacements of the band maximum in the region 2.0–2.17 eV and the absorption on the long-wavelength shoulder of the band can be attributed to the different size of the particles and their possible surface oxidation in an environment consisting of variable-valence oxides. On the basis of the photolytic model of photochromism the similarity of the ΔD curves for different irradiation doses indicates that an increase in the number of Cu_n particles predominates over an increase in particle size.

Depending on the thermal history of the FKhS-7 samples, a band with a maximum at 2.53–2.55 eV was also observed in the spectra ΔD.[2] According to the shell model of the formation of photolytic Cu_n this band can be assumed to be of a molecular-center nature, i.e., the formation of Cu_2 or Cu_3 type centers in a near-surface region of the light-sensitive CuCl phase, enriched with mobile copper,[3] under the conditions of photoinduced redistribution of the electron density. Considering the sizes of the CuCl particles and the presence of a phase boundary (PB) with the matrix, photogenerated equilibrium Cl–O centers, for example, Cl_2O,[4] which contribute to the total spectrum ΔD, must also be admitted.

In the present paper we attempt to demonstrate that copper oxides (especially cuprous oxide) play an important role in the photochroism of copper-halide glasses and to elucidate the mechanisms of the corresponding photochemical reactions. The analysis of the photochromic absorption spectra is based on a study of the absorption spectra of copper oxides obtained in various matrices. Specifically, the dominant band in the absorption spectrum of semitransparent mirror layer of copper thermally oxidized in a dry-air atmosphere (Fig. 1, A) is the band at 2.53 eV, whose long-wavelength shoulder contains structure in the region 2.38–2.43 eV. This band is also dominant in the absorption spectrum of the cathode deposit of copper oxide obtained by the standard electrochemical technology (Fig. 1, B). Its half-width is characteristic for the analogous band in the induced absorption spectrum of FKhS-7 glass. Here it should be noted that the nature of the

long-wavelength bands in the region 1.9–2.17 is a subject of debate.[5,6] Selective irradiation of FKhS samples does not completely elucidate the nature of photochromism.[7,8]

Figure 1, C shows the changes in the spectrum ΔD of FKhS-7 glass under successive equal-energy selective irradiation in the spectral regions 2.45–2.75 eV (curve *1*) and 1.75–2.25 eV (curve *2*). In the first case irradiation produces a 40 meV shift of the band maximum into the short-wavelength region, it broadens and increases the intensity of the band, and it increases the intensity of the absorption in the blue-violet region of the spectrum. The intensity of the 2.55 eV band remains unchanged. Under subsequent irradiation in the yellow-red region of the spectrum the intensity of this band increases, the absorption in the violet region decreases negligibly, and the half-width of the long-wavelength band decreases and its maximum undergoes an approximately 20 meV bathochromic shift. These results undoubtedly show that multiple processes are responsible for photochromism.

The strong temperature dependence of the spectrum ΔD in the temperature range 250–300 K ordinarily used in prac-
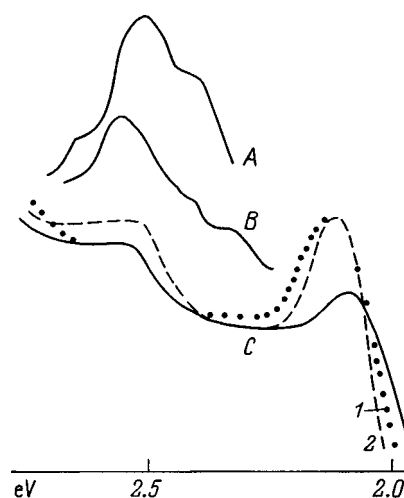

FIG. 1. Absorption spectra of a thermally oxidized film of electrolytic copper (A) and a red cathode deposit of copper oxide (B) as well as the spectral variations of the photoinduced absorption of FKhS-7 glass (C, solid curve) after additional illumination with blue (*1*) and red(*2*) light.

tical applications, specifically, the nonlinear growth of ΔD with decreasing temperature, which is also characteristic for silver-halide FKhS glasses,[9] does not agree with the increase in the efficiency of the photolytic process accompanying a temperature increase in ion mobility in AgHal and CuHal. By changing the composition of the initial charge glasses with a CuCl phase can be used as FKhS glasses or as filters which cut off UV radiation.[10] Therefore the efficiency of photochromism is determined not only by the presence of a light-sensitive phase. The direct dependence of the melting temperature of this phase on its average size,[11] the short response time of photochromism, and the dependence of the low-energy limit of photoactive photons on the history of the sample[2] definitely indicate that the phase boundary (PB) (matrix–light-sensitive phase) influences the efficiency of photochromism. The fact that photoinduced processes and PBs are associated with one another has been shown previously on model objects with AgI and CuI nanocrystals (NCs).[12,13] Specifically, films with AgI and CuI NCs suspended in cavities in a gelatin matrix are photographically stable and can be used as filters for cutting off UV radiation, but after thermal adsorption of gelatin on the surface of NCs the same samples become photographically effective. In this context the photolytic precipitation of colloidal and precolloidal Cu_n particles on PBs (just as Cl) on the side of the free CuCl surface is unlikely because of spatial confinement.

A band similar to the high-energy section of the visible spectrum of FKhS-7 with thermally-induced yellow coloration is observed in the difference absorption spectra of FKhS-7 samples (Fig. 2, A) in the spectral region 2.75–3.17 eV. With the exception of the 2.97–3.01 eV band the ''yellowness'' spectrum and the spectrum of excitons, which are observed in a glass matrix in Ref. 14, correlate, so that this spectrum can be attributed to a cuprous oxide phase in contact with CuHal.[15,16]

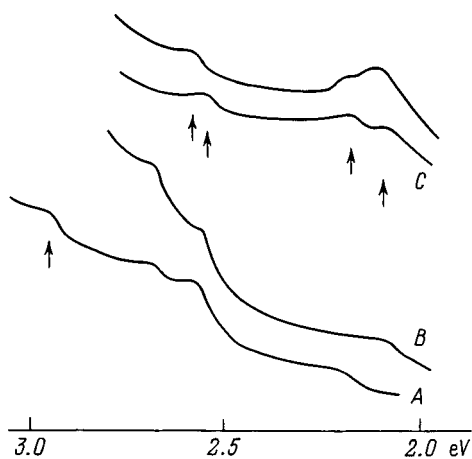Besides the bands characteristic for cuprous oxide, bands with maxima at 2.97 and 2.85 eV are also observed in the absorption spectra of the light-yellow commercial copper oxide glasses subjected to secondary low-temperature heat treatment.[5] Moreover, the 2.97–3.04 eV band is dominant in the spectrum of an aqueous ammonia solution of cuprous oxide.[4] Assuming that the mobile copper diffuses toward PBs as regions of distortion of the uniformity of the electrical potential,[3,13] the FKhS absorption bands at 2.85–3.0 eV can be attributed to copper-oxide centers.[17] Therefore a copper-oxide nature of both the FKhS-7 absorption edge and the spectrum ΔD as well as the formation of a solid solution on PBs (matrix–CuCl) as a result of the concentration of mobile copper and the associated Coulomb redistribution of the electron density are admissible. The possibility of photo-induction of a solid-phase reaction is demonstrated in Refs. 8 and 13. At the same time photoirradiation can also lead to phase separation of the solid solution.[18] The change in the direction of the solid-phase chemical reaction and the composition of the product obtained could depend strongly on the temperature of the sample and the spectral composition of the exciting light.[17]

Figure 2, B shows the spectrum ΔD of a FKhS-7 samples irradiated in the 3.4 eV band. Here the 2.72 and 2.58 eV bands characteristic for Cu_2O (Ref. 14) are resolved. The absorption spectrum of the other sample was later altered by irradiating with 2.75 eV light (Fig. 2, C), which resulted in a hypochromic shift of the 2.54 eV band by 40 meV and a relative increase in the intensity of the 2.1 eV band against the background of the 2.15 eV band, which, just as the 2.54–2.58 eV band, cannot be associated with excitonic absorption in Cu_2O.[14]

The dark relaxation spectra D for the FKhS-7 sample preirradiated with sunlight at 255 K, which were obtained in a regime of heating up to 280 K, are presented in Fig. 3, A. Figure 3, B shows the spectra ΔD for a FKhS-7 sample with weak light-yellow coloration. These spectra were recorded immediately after irradiation with sunlight at 285 K and 30 min later. On the basis of the copper and copper-oxide nature of ΔD the 1.9–1.96 eV band corresponds to Cu^{2+}–Re centers, whose thermodynamic formation is possible on PBs with Cu^+ interacting with a dissociating Cl–O center, since the oxidation $Cu^+ \rightarrow Cu^{2+}$ by chlorine on a free surface is possible only in the presence of a copper deficiency, which contradicts the copper-colloidal nature of the 2.1 eV band. In Ref. 1 the 2.38–2.48 eV band is attributed to molecular-size copper centers.

The resonance character of the excitation of the bands in the spectrum ΔD makes it possible to give an energy estimate of the 2.38–2.48 eV band on the basis of the proposed model of centers.[19] The binding energies for Cu_2 and Cu_3 are, respectively, 1.98 and 2.97 eV,[1] while the binding energy of the dynamically equilibrium center can be taken as the average of these values, which is 2.475 eV. The affinity of Cu_2^+ can be estimated, by analogy with the same quantity for Ag_2^+,[19] as $1.5E_a$ (Cu^0), which is 1.845 eV. The binding energy of a Cu_3 center in a dynamically dissociated state can be estimated as the average of 1.845 and 2.97 eV, i.e., 2.4 eV. We note that the average of 2.97 and 1.23 eV[4] is 2.1 eV, so that Cu_n on a CuCl surface can be regarded as a superposition of interacting Cu_3 centers, which satisfies the model of

FIG. 2. Thermally induced ''yellowness'' spectra in FKhS-7 glass (A) and photoinduced absorption in FKhS-7 glass under irradiation in the 3.4 eV band (B) and of an another FKhS-7 sample under irradiation in the 3.4 eV band (C, lower curve) followed by irradiation in the 2.75 eV band (C, upper curve). All spectra were recorded relative to the unirradiated sample. The ''yellowness'' spectrum was recorded relative to the colorless section.
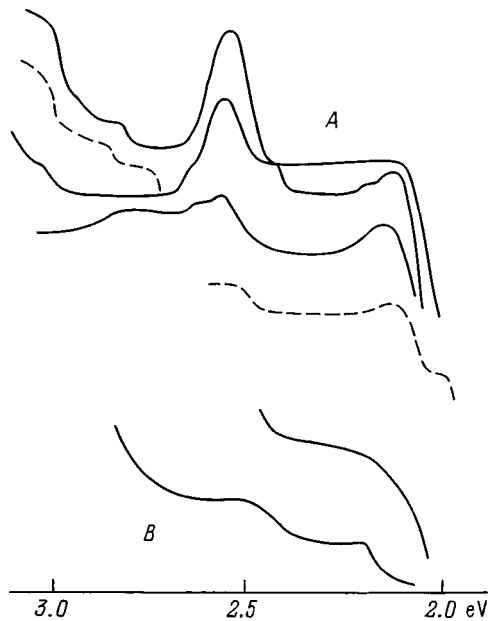
FIG. 3. Comparative spectra of dark relaxation of photoinduced absorption. The spectra were recorded with the sample heated from 255 to 280 K (A): dashed curves — individual sections of the spectrum which were obtained with repeated irradiation–relaxation cycles. Sectral variations of the photo-induced absorption for the FKhS-7 sample with a weak light-yellow color (B). All spectra were recorded with respect to the unirradiated samples.

phase separation of the solid solution CuCl(Cu).[20]

An increase of ΔD under irradiation is accompanied by a decrease of the intensity and an increase of the half-width of the 3.2–3.21 eV band, associated with excitonic absorption in CuCl,[2,7,10,11] as well as by a shift of the band maximum to 3.24–3.25 eV. As ΔD relaxes, the spectral characteristics of the indicated band are restored. However, incomplete relaxation of ΔD in the spectral region of the Cu_n band and the associated incomplete restoration of the spectral characteristics of the 3.2 eV band are observed in a number of cases. Similar changes for this band have been observed in CuCl–CdCl$_2$ films.[21] The possibility of relaxation of ΔD presupposes that a thin layer of CuCl nanocrystals near PBs participates in the photoinduced solid-phase reaction. In this connection the short-wavelength shift of the 3.2 eV band could be due to an increase in the influence of the structure of the substrate with a cuprous oxide phase formed[12] and a change in the relative copper–halogen concentration.[3]

To determine more accurately the role of the ion stage in the mechanism of photochromism, curves of the photoinduced charge currents and of the dark discharge of a capacitor cell with FKhS-7 samples were obtained. The nearly exponential character of the current curves is similar to the curves of the kinetics of the growth and relaxation of ΔD.[2] This indicates ion diffusion and corresponds to a solid-phase chemical reaction. At the same time the photocurrent curves obtained in a differential detection regime[22] in a number of successive traces (irradiation–dark relaxation) show oscillations against the background of a slow current component. In the opinion of one of the authors of Ref. 22 this corresponds to Coulomb separation of charged particles and an associated

compensatory motion of the charge carriers. Such a process is possible only on a free CuCl surface and corresponds to segregation of Cu_n particles as a result of their resonance excitation and ionization. In the model of photochromism as a reversible solid-phase reaction the formation of Cu_n on a free CuCl surface is a consequence of an increase in the concentration of mobile copper (Cu^+) and its Coulomb displacement away from the PBs (CuCl–matrix) as a result of an increase in the concentration of h^+ centers as well as mobile Na^+ ions.[23] Conversely, in the process of thermal synthesis of the phase Cu_2O at 300–350 °C in the near-surface region of CuCl the binding of mobile Cu^+ from CuCl will increase the Na^+ concentration in CuCl (segregation of impurities into a melt zone). As a result of the relatively high dissociation energy of cuprous oxide and the absence of a weakly bound or free oxygen, low-energy excitation of the heterosystem CuCl(NA)–Cu_2O does not lead to reduction of cuprous oxide. The possibility of restoration of photochromic processes after the yellow-colored samples are heated at 450–500 °C means that the system CuCl–Cu_2O–alkali oxides is glass-forming. This follows from, among other things, the synthesis of ion-selective glasses with a high percentage of CuHal or Cu_2O.

It follows on this basis that in FKhS-7 glass a heterosystem treated in photomaterials as a substrate-ion-selective epitaxy structure,[13,18,22] where the epitaxy together with PBs are regions of accumulation of photoinduced and mobile centers, exhibits photochromic properties. In Ref. 13 it is shown that depending on the conditions at the PBs and the chemical composition of the constituents of the heterosystem the same substance can act as a substrate or as epitaxy. In this connection, considering the polymorphism of CuHal (specifically, accompanying a change in a copper–halogen ratio[3]), the reversibility of the solid-phase photochemical reaction, as well as the constancy of the copper–oxygen ratio in the region of the easy-melting phase around CuCl, Cu_2O can be assumed to be polymorphic.[24] The change in the slope of the temperature dependence of the electrical conductivity of Cu_2O films at 350 °C (Ref. 25) can also be regarded as an indication of structural changes, for example, transfer of some cations into interstices and formation of a $Cu_3O_2(Cu_2)$ structure as an intermediate phase between Cu_2O and CuO.

A sharp decrease in ΔD in the temperature range 250–290 °C was observed in Ref. 26.

Here it is worthwhile noting that after FKhS-7 is heated at 300 °C a light-yellow coloration with ΔD characteristic for yellow Cu_2O is observed. In the case of a structural phase transition in a CuCl nanocrystal, which is possible in the indicated temperature range, a long-wavelength shift of the absorption edge and an increase of ΔD are most likely.[3] The thermally induced decrease of ΔD is most likely due to the removal of intrinsic impurities or dopants from CuCl in the case of amorphization or a structural phase transition in the substrate material.[27] The reversible restoration of the temperature dependence of the increase in ΔD, starting at 300 °C (Ref. 26) supports a structural phase transition, where the concentration of cationic vacancies increases with the symmetry of the crystal lattice (principle of an ion-selective electrode). In this connection we examined the tem-

perature dependence of the spectral variations of a yellow film of cuprous oxide on a glass surface with an electrically conducting SnO_2-based coating obtained by cathode deposition in an electrochemical cell; this system can be viewed as a model system for the Cu_2O phase in FKhS-7. It was shown that a decrease of absorption in the region 2.75–3.1 eV and a proportional increase of absorption in the region 1.9–2.63 eV are observed in the same temperature range. These changes are reversible under oxygen deficiency conditions.

In summary, photo- or thermally-induced displacement of the equilibrium and a change in the direction of the solid-phase chemical reaction are possible in the heterosystem CuCl–Cu_2O—matrix as a result of the polymorphism of both compounds.

[1] Yu. I. Petrov, *The Physics of Small Particles* [in Russian], 1982, 389 pp.

[2] A. L. Ashkalunin, P. M. Valov, V. I. Leĭman, and V. A. Tsekhomskiĭ, Fiz. Khim. Stekla **10**, 325 (1995).

[3] V. A. Voll, Zh. Tekh. Fiz. **65**(10), 191 (1995) [Tech. Phys. **40**, 1080 (1995)].

[4] A. I. Efimov (ed.), *The Properties of Inorganic Compounds* [in Russian] Khimiya, Moscow (1983), 390 pp.

[5] I. Kotsik, I. Nebrzhenskiĭ, and I. Fanderlik, *The Coloration of Glass* [Stroĭizdat, Moscow (1983), 211 pp.].

[6] D. G. Galimov, A. M. Gubaĭdullina, and A. I. Neĭch, Fiz. Khim. Stekla **13**, 50 (1987).

[7] L. B. Glebov, N. V. Nikonorov, and G. T. Petrovskiĭ, Dokl. Akad. Nauk SSSR **280**, 1110 (1985) [Sov. Phys. Dokl. **30**, 147 (1985)].

[8] V. M. Marchenko, Fiz. Khim. Stekla **21**, 359 (1995).

[9] L. V. Gracheva and V. A. Tsekhomskiĭ, Opt. Mekh. Promst., **44**(9), 29 (1977) [Sov. J. Opt. Technol. **44**, 540 (1977).

[10] V. I. Vasil'ev, Candidate's Dissertation [in Russian], St. Petersburg (1996), 155 pp.

[11] A. A. Oshchenko, Candidate's Dissertation [in Russian], Leningrad (1989), 128 pp.

[12] A. L. Kartuzhanskiĭ, L. K. Kudryashova, V. A. Reznikov *et al.*, Opt. Spektrosk. **69**, 1323 (1990) [Opt. Spectrosc. **69**, 784 (1990)].

[13] V. A. Voll, A. V. Barmasov, and A. V. Struts, Neorg. Mater. **30**, 841 (1994).

[14] E. F. Gross, Chang Kuang-yin, and L. E. Solov'ev, Dokl. Akad. Nauk SSSR **146**, 577 (1962) [Sov. Phys. Dokl. **7**, 821 (1963)].

[15] E. E. Podorova and V. A. Tsekhomskiĭ, Fiz. Khim. Stekla **16**, 555 (1990).

[16] V. V. Golubkov and A. A. Onushchenko, Fiz. Khim. Stekla **23**, 389 (1997).

[17] V. B. Aleskovskiĭ, *The Chemistry of Supermolecular Compounds* [in Russian], St. Petersburg University (1996), 253 pp.

[18] A. V. Barmasov and V. A. Reznikov, "Crystallization in the solid phase," Dep. VINITI No. 6859-V89 [in Russian], All-Union Institute for Scientific and Technical Information, Moscow (1989), 29 pp.

[19] V. A. Voll, Fiz. Tekh. Poluprovodn. **29**, 2071 (1995) [Semiconductors **29**, 1081 (1995)].

[20] V. V. Slezov and V. V. Sagalovich, Usp. Fiz. Nauk **151**, 67 (1987) [Sov. Phys. Usp. **30**, 23 (1987)].

[21] O. N. Yunakova, V. K. Miloslavskiĭ, and Bayarma Guan-Adzhavan, Zh. Nauchn. Prikl. Fotogr. Kinematogr. **35**(5), 369 (1990).

[22] V. A. Voll and O. A. Trofimov, Zh. Fiz. Khim. **2**, 323 (1996).

[23] V. V. Golubkov and V. A. Tsekhomskiĭ, Fiz. Khim. Stekla **12**, 206 (1986).

[24] K. Sugasaka, A. Fujii, S. Kato, and D. Mijuguti, RZhKhim., No. 5B785 (1970).

[25] R. E. Krzhizhanovskiĭ and Z. Yu. Shtern, *The Thermophysical Properties of Nonmetallic Materials* [in Russian], Énergiya, Leningrad (1973), 333 pp.

[26] S. B. Gorbatova, A. M. Zyabnev, and S. L. Kraevskiĭ, Fiz. Khim. Stekla **19**, 266 (1993).

[27] A. L. Kartuzhanskiĭ, L. K. Kudryashova, A. V. Barmasov, and V. A. Reznikov, Opt. Spektrosk. **66**, 332 (1989) [Opt. Spectrosc. **66**, 191 (1989)].

Propagation of a monochromatic electromagnetic plane wave in a medium with nonsimple motion

V. O. Gladyshev

Egor'evsk College of Civilian Aviation Technology, 140303 Egor'evsk, Moscow Region, Russia
(Submitted November 11, 1996; resubmitted October 28, 1997)
Zh. Tekh. Fiz. **69**, 97–100 (May 1999)

An exact analytical solution is obtained for the trajectory of the wave vector of a monochromatic
electromagnetic plane wave in a medium with nonsimple motion. It is shown that the
spatial dragging of the electromagnetic wave by the moving medium can be described correctly
in the general case only if relativistic terms of order β^2 are taken into account. © *1999*
American Institute of Physics. [S1063-7842(99)01705-5]

INTRODUCTION

The solution of the dispersion relation for the propagation of a monochromatic electromagnetic plane wave in a moving medium reduces to finding the wave vector $\mathbf{k}_2$ of the electromagnetic wave in the medium. For propagation of electromagnetic radiation in a medium with nonsimple motion it is necessary to obtain a solution for each local region along the entire propagation trajectory of the wave, since the wave vector at each point of the trajectory depends on the velocity vector $\mathbf{u}_2$ of the medium.[1]

This dependence is a consequence of the fact that the Fizeau effect is a particular case of the spatial dragging of the light by a moving medium[2] and can be investigated experimentally. For example, in experiments on laser ranging of a space vehicle (SV) it was found that the Fizeau effect has an appreciable influence on the direction of the laser beam passing through a moving quartz reflector.[3,4] Since the effect should appear in interferometric measurements at comparatively low velocities of the propagation medium of the electromagnetic wave, it should affect the results of a wide class of experiments.

The investigation of this phenomenon is also of interest from a different standpoint, since it can be viewed as a precise test of the electrodynamics of moving media in which the interaction of an electromagnetic wave with the moving matter can be analyzed on an atomic scale on the basis of macroscopic observations. For this reason there naturally arises the problem of finding analytical equations describing the trajectory of the wave vector in a medium with nonsimple (nonrectilinear) motion.

We note that in the literature the description of the effect of the motion of the medium on the propagation of electromagnetic radiation is ordinarily limited to approximate calculations. This is admissible in calculations of the Fizeau effect, which can be described quite accurately taking into account only the first-order relativistic terms. However, this simplification is unacceptable for subtle phenomenon such as the curvature of the trajectory of an electromagnetic wave, whose magnitude is determined by second-order relativistic terms.

FORMULATION OF THE PROBLEM

Let us consider an inertial coordinate system in which a medium with permittivity ε_1 and magnetic permeability μ_1 is at rest in the half space $Z<0$ and a medium with ε_2 and μ_2, which are measured in its own coordinate system, moves with arbitrary velocity $\mathbf{u}_2$ in the half space $Z>0$. A tangential discontinuity of the velocity exists at the interface between the media. We assume that the velocity field is constant along the Y axis.

Neglecting the dispersion of the moving medium and taking $\varepsilon_1\mu_1=1$ the wave vector $\mathbf{k}_2$ in the second medium can be expressed as[5]

$$k_{2x}=\frac{\omega_0}{c}\sin\vartheta_0, \tag{1a}$$

$$k_{2z}=\frac{\omega_0}{c}\frac{1}{1-\kappa_2^2\beta_{2z}^2\gamma_2^2}\{-\kappa_2\beta_{2z}\gamma_2^2(1-\beta_{2x}\sin\vartheta_0)$$

$$\pm(\cos^2\vartheta_0(1-\kappa_2^2\beta_{2z}^2\gamma_2^2)+\kappa_2\gamma_2^2(1-\beta_{2x}\sin\vartheta_0)^2)^{1/2}\}, \tag{1b}$$

where $\gamma_2^{-2}=1-(\beta_{2z}^2+\beta_{2x}^2)$, $\kappa_2=\varepsilon_2\mu_2-1$, $\beta_{2x}=u_{2x}/c$, $\beta_{2z}=u_{2z}/c$, $\mathbf{u}_2$ is the velocity vector of the second medium, ϑ_0 is the angle of incidence of the electromagnetic wave on the interface of the two media, ω_0 is the angular frequency of the electromagnetic wave, c is the speed of light in vacuum, and the sign is chosen so that the wave propagates away from the interface between the two media.

The form of the expression for k_{2z} imposes a restriction on the function $\beta_2(x,z)$ for which analytical solutions of the equation of the trajectory of the wave vector in the medium exist.

Using the function (1b) in searching for a solution of the form $z=f(x)$ leads to an implicit integral equation $z=\int_0^{x_{\max}(x,z)}f(x,z)dx$, which does not have an analytical solution in the general case.[6] However, there is an analytically solvable case where the spatial character of the dragging of the light is most naturally manifested.

Let the velocity u_2 be a function of the coordinates x and z of the form

$$\beta_2^2 = \frac{\omega^2}{c^2}(R_0 - z)^2 + \frac{\omega^2}{c^2}x^2, \tag{2}$$

which corresponds to rotation with angular velocity ω relative to the center $(0, R_0)$. This function determines the parameters $u_{2x} = \omega(R_0 - z)$ and $u_{2z} = \omega x$ as functions of the independent coordinates. The function (2) with both components requires the use of numerical methods, which was done in Ref. 2, since analytical methods lead to computational errors, due to the dropping of terms in the series expansion, and so on, that are larger than the effect under investigation.

On the other hand the tangential component of the velocity of the medium influences the spatial character of the dragging of the light (the curvature of the trajectory of the wave vector). For this reason this is the most interesting case for studying the spatial dragging of light by a moving medium.

Let only a tangential component β_{2x} be present in the moving medium and $\beta_{2z} = 0$. This corresponds to a shear flow with velocity varying linearly with distance from the boundary. Then the angle of refraction ϑ_2 can be written as

$$\tan^2 \vartheta_2(z) = \left(\frac{k_{2x}}{k_{2z}}\right)^2$$

$$= \frac{\alpha^2(1 - \beta_x^2(z))}{\gamma^2(1 - \beta_x^2(z)) + (n_2^2 - 1)(1 - \alpha\beta_x(z))^2}, \tag{3}$$

where $\alpha = \sin\vartheta_0$, $\gamma = \cos\vartheta_0$, $n_2 = \sqrt{\varepsilon_2\mu_2}$, and $\beta_x(z) = u_{2x}(z)/c$.

The wave vector $k_0 = 2\pi/\lambda_0$ of the incident electromagnetic wave satisfies the relation $k_0 \gg 1/R_0$, which makes it possible to use the solution of the wave equation for an electromagnetic plane wave undergoing a tangential velocity discontinuity at the interface between the media[4] for each local region of the medium.

We shall be interested in the equation describing the trajectory of $\mathbf{k}_2$, i.e., the analytical dependence $x = f(z)$. Evidently, the difference relation

$$\Delta x_i = \tan\vartheta_2(z_i)\Delta z_i,$$

where $\Delta x_i = x_i - x_{i-1}$, $\Delta z_i = z_i - z_{i-1}$, $z_i = \Sigma_{j=1}^{i}\Delta z_j$, and $x_i = \Sigma_{j=1}^{i}\Delta x_j$, holds for each local region of the medium.

Then a relation between the instantaneous coordinates can be obtained by summing and switching to an integral in the limit $\Delta z_i \to 0$

$$x = \int_0^z \tan\vartheta_2(z)dz. \tag{4}$$

A characteristic feature of this expression is that the limits can be set arbitrarily, but we do not have precise information about the point where the trajectory intersects, for example, a prescribed cylindrical surface of radius R_0. Therefore, generally speaking, the upper limit of the integral is variable. We also note that the expression for the angle of refraction is exact and contains quadratic terms. This is of fundamental significance for studying the spatial dragging of light by a moving medium.

ANALYTICAL SOLUTION OF THE EQUATION FOR THE TRAJECTORY OF THE WAVE VECTOR OF AN ELECTROMAGNETIC WAVE

We shall seek the solution of Eq. (4) in a general form. For this, we substitute expression (3) and switch to the new variable β_x. After transformations we obtain

$$x = \tau \int_{\beta_1}^{\beta_2} \frac{\beta_2(\beta_x - 1)d\beta_x}{\sqrt{G^4(\beta_x)}}, \tag{5}$$

where

$$\tau = \frac{c}{\omega}\frac{\alpha}{\sqrt{1 - n_2^2\alpha^2}},$$

$$G^4(\beta_x) = (a - \beta_x)(b - \beta_x)(\beta_x - c)(\beta_x - d),$$

$$\beta_{x1,2} = \frac{\alpha(1 - n_2^2) \pm \gamma^2 n_2}{1 - n_2^2\alpha^2},$$

$$a = \beta_{x1}, \quad b = -c = 1, \quad d = \beta_{x2}.$$

The expression contains the root of a quartic polynomial and it can be shown that Eq. (5) can be represented as a composition of elliptic integrals. The integration limits are determined from the expression for $\beta_{2x}(z)$ for the initial and final coordinates z_1 and z_2 of the trajectory of the wave vector. We introduce the notation

$$J_S = \int \frac{d\beta_x}{(\beta_x - 1)^S \sqrt{G^4(\beta_x)}}. \tag{6}$$

Then the coordinate x can be expressed as

$$x = \tau(J_{-2} - 2J_{-1})\big|_{\beta_1}^{\beta_2}. \tag{7}$$

In order to reduce J_{-2} to tabulated integrals, its order must be increased. Let us expand $G^4(\beta_x)$ in powers of $(\beta_x - 1)$

$$G^4(\beta_x) = b_0(\beta_x - 1)^4 + b_1(\beta_x - 1)^3$$
$$+ b_2(\beta_x - 1)^2 + b_3(\beta_x - 1) + b_4,$$

where

$$b_0 = 1, \quad b_1 = 4 - (\beta_{x1} + \beta_{x2}),$$

$$b_2 = 5 + \beta_x\beta_{x2} - 3(\beta_{x1} + \beta_{x2}),$$

$$b_3 = 2(1 + \beta_{x1}\beta_{x2} - (\beta_{x1} + \beta_{x2})),$$

and $b_4 = 0$. Integrating the first derivative of the product

$$\sqrt{G^4(\beta_x)}(\beta_x - 1)^{-S},$$

we obtain a recurrence relation that makes it possible to lower the order of the elliptic integral

$$b_0(2 - S)J_{S-3} + \frac{b_1}{2}(3 - 2S)J_{S-2} + b_2(1 - S)J_{S-1}$$

$$+ \frac{b_3}{2}(1 - 2S)J_S - b_4 SJ_{S+1} = \sqrt{G^4(\beta_x)}(\beta_x - 1)^{-S},$$

$$S = 1, 2, 3, \ldots. \tag{8}$$

Using Eq. (8) with $S=1$ and $b_4=0$ we obtain an expression for J_{-2}

$$J_{-2}=\frac{1}{2b_0}\left(\frac{2\sqrt{G^4(\beta_x)}}{\beta_x-1}-b_1J_{-1}+b_3J_1\right). \quad (9)$$

Substituting expression (9) into Eq. (7), the expression for x becomes

$$x=\frac{\tau}{2b_0}\left(\frac{2\sqrt{G^4(\beta_x)}}{\beta_x-1}+b_3J_1+(4b_0-b_1)J_{-1}\right)\Bigg|_{\beta_1}^{\beta_2}. \quad (10)$$

We note that $a>b\geq\beta_x>c>d$, and we introduce the notations

$$I_1=\int_c^{\beta_x}\frac{\beta_xd\beta_x}{\sqrt{G^4(\beta_x)}}, \quad I_2=\int_c^{\beta_x}\frac{d\beta_x}{\sqrt{G^4(\beta_x)}}, \quad I_3=-J_1|_c^{\beta_x}.$$

Then Eq. (10) can be expressed in terms of tabulated integrals

$$x=\frac{\tau}{2b_0}\left(\frac{2\sqrt{G^4(\beta_x)}}{\beta_x-1}-b_3I_3+(4b_0-b_1)(I_1-I_2)\right)\Bigg|_{\beta_1}^{\beta_2},$$

$$I_1=2g[(c-d)\Pi(\varphi,n_1,k)+dF(\varphi,k)],$$

$$I_2=2gF(\varphi,k),$$

$$I_3=2gq[(c-d)\Pi(\varphi,n_2,k)+(1-c)F(\varphi,k)],$$

$$g=\frac{1}{\sqrt{(a-c)(b-d)}}, \quad q=\frac{1}{(1-c)(1-d)}, \quad (11)$$

where $F(\varphi,k)$ and $\Pi(\varphi,n_i,k)$ are normal elliptic integrals of the first and third kinds, to which correspond the characteristics n_1 and n_2, the amplitude φ, and the modulus k given by

$$n_1=\frac{b-c}{b-d}, \quad n_2=\frac{(b-c)(1-d)}{(b-d)(1-c)},$$

$$\varphi=\sin^{-1}\sqrt{\frac{(b-d)(\beta_x-c)}{(b-c)(\beta_x-d)}}, \quad k=\sqrt{\frac{(b-c)(a-d)}{(a-c)(b-d)}}.$$

Finally, substituting the coefficients a, b, c, and d we obtain

$$x=\tau\Bigg(c_1\Pi(\varphi,n_2,k)-c_2\Pi(\varphi,n_1,k)$$

$$-c_3F(\varphi,k)-\frac{\sqrt{G^4(\beta_x)}}{1-\beta_x}\Bigg)\Bigg|_{\beta_1}^{\beta_2}, \quad (12)$$

where

$$c_1=\frac{1}{t}(1+\beta_{x2})(1-\beta_{x1}),$$

$$c_2=\frac{1}{t}(1+\beta_{x2})(\beta_{x2}+\beta_{x1}),$$

$$c_3=\frac{1}{t}[2(1-\beta_{x1})+(1-\beta_{x2})(\beta_{x2}+\beta_{x1})],$$

$$t=\sqrt{(1+\beta_{x1})(1-\beta_{x2})}.$$

Comparing the computational results obtained with Eq. (12), using tables of elliptic integrals,[7] with the results of direct numerical calculations using Eq. (4) shows that the accuracy of the analytical calculations depends on the resolution of the tables of elliptic integrals and that interpolation must be used. Nonetheless, the expression obtained is exact, and it is desirable to construct more accurate tables of elliptic integrals in the parameter range corresponding to the experimental data.

It is also evident that $n_2=1$ for any d. In this case $\Pi(\varphi,n_2,k)$ can be expressed in terms of elliptic integrals of the first and second kinds:

$$\Pi(\varphi,n_2=1,k)=F(\varphi,k)-\sec^2\alpha(E(\varphi,k)-\tan\varphi\Delta\varphi),$$

$$\Delta\varphi=\sqrt{1-k^2\sin^2\varphi}, \quad k=\sin\alpha.$$

The use of this expression decreases the interpolation error because the tables of $F(\varphi,k)$ and $E(\varphi,k)$ is more accurate.

ESTIMATION OF THE EFFECT OF TERMS OF ORDER β^2

Here it is appropriate to raise the question of whether or not terms of order β^2 are needed to describe the three-dimensional Fizeau effect. Indeed, the electrodynamics formulas for the longitudinal Fizeau effect are linear in the velocity of the medium and terms of order β^2 are negligible, though they are present in the exact solution of the dispersion relation. However, for the three-dimensional Fizeau effect, i.e., when a transverse component of dragging of the wave appears, it may be necessary to take account of the terms with β^2. To justify this assertion we shall obtain an expression for the path length of the light beam in a medium with a chosen law of motion neglecting β^2. In the geometric-optics approximation the path length of the light beam can be written as

$$S=\int_0^z\sqrt{1+\tan^2\vartheta_2(z)}dz, \quad (13)$$

where

$$\tan^2\vartheta_2\cong\frac{\alpha^2}{\gamma^2+(n_2^2-1)(1-2\alpha\beta_x)}. \quad (3')$$

Changing to the new variable and performing the transformation, we obtain

$$S=-\frac{c}{\omega}\int_{\beta_1}^{\beta_2}\sqrt{\frac{a-\beta_x}{b-\beta_x}}d\beta_x, \quad (14)$$

where

$$a=\frac{n_2^2}{2\alpha(n_2^2-1)}, \quad b=\frac{n_2^2-\alpha^2}{2\alpha(n_2^2-1)}.$$

Performing the integral in Eq. (14) gives

$$S=\frac{c(a-b)}{\omega}\left(\frac{1}{2}\ln\left|\frac{+1}{-1}\right|+\frac{\psi}{\psi^2-1}\right)\Bigg|_1^2, \quad (15)$$

where

$$\psi = \sqrt{\frac{a - \beta_x}{b - \beta_x}}, \quad \psi_{1,2} = \sqrt{\frac{a - \beta_{x1,2}}{b - \beta_{x1,2}}}.$$

For $\omega = 0$ the length of the trajectory up to intersection with the straight line $z = R_0$ can be written as $S_0 = R_0 / \cos\vartheta_2$, where ϑ_2 can be found from Snell's law.

Then the difference of the path length S, taking account of only terms of order β, and the path length S_0 without rotation will approximately characterize the magnitude of the Fizeau effect in a moving medium.

Numerical calculations show that the curvature of the propagation trajectory of an electromagnetic wave in a medium with rotation[6] is of the same order of smallness as the error in the calculations performed using Eq. (15). This confirms that β^2 must be taken into account in order to describe correctly the spatial dragging of electromagnetic radiation by a moving medium.

CONCLUSIONS

The Fizeau effect is usually characterized by the magnitude of the drift of the phase velocity of the superposition field of the excitation and secondary electromagnetic waves in a moving medium. It is also convenient to use as the characteristic of the longitudinal dragging of the electromagnetic wave the phase difference between the waves which have passed through the moving medium in opposite directions. An additional effect appears in the case of three-dimensional dragging of an electromagnetic wave — the deflection of the trajectory of the wave vector of the superposition wave in the medium. The analytical solutions of the equation for the trajectory of the wave vector in a moving medium can be used to describe this phenomenon.

In closing, we note that the solution of Eq. (4) can be represented as a composition of elliptic integrals not only for the chosen law of motion of the medium. This opens up the possibility of using analytical methods to describe the trajectory of a wave vector in media with more complicated laws of motion.

[1] S. Solimeno, B. Coroziniani, and P. DiPorto, *Guiding, Diffraction, and Confinement of Optical Radiation* [Academic Press, New York, 1986; Mir, Moscow, 1989, 664 pp.].

[2] V. O. Gladyshev, JETP Lett. **58**, 569 (1993).

[3] V. P. Vasil'ev, V. A. Grishmanovskiĭ, L. F. Pliev, and T. P. Startsev, JETP Lett. **55**, 316 (1992).

[4] V. P. Vasil'ev, L. I. Gusev, J. J. Dengan, and V. D. Shargorodskiĭ, Radiotekh., No. 4, 80 (1966).

[5] B. M. Bolotovskiĭ and S. N. Stolyarov, Usp. Fiz. Nauk **159**, 155 (1989) [Sov. Phys. Usp. **32**, 813 (1989)].

[6] V. O. Gladyshev, Pis'ma Zh. Tekh. Fiz. **19**(19), 23 (1993) [Tech. Phys. Lett. **19**(10), 611 (1993)].

[7] M. Abramowitz and I. Stegun (Eds.), *Handbook of Mathematical Functions* [Dover, New York, 1965; Nauka, Moscow, 1979, 832 pp.].

Translated by M. E. Alferieff

Branching of stationary inversion states in a cavity-pumped paramagnetic maser amplifier

D. N. Makovetskiĭ, A. A. Lavrinovich, and N. T. Cherpak

Institute of Radiophysics and Electronics, National Academy of Sciences of Ukraine,
310085 Kharkov, Ukraine
(Submitted January 14, 1998)
Zh. Tekh. Fiz. **69**, 101–105 (May 1999)

The nonlinear dynamics in a paramagnetic maser amplifier is investigated experimentally and theoretically under conditions such that the active medium exerts a sufficiently strong feedback influence on the pump field to induce population inversion of the spin levels. Branching of the inversion ratio due to the onset of bistability in the nonlinear microwave cavity of the pump at a frequency of 150 GHz is observed experimentally. Conditions are determined for the possible excitation of new nonlinear resonances when the spin system is inverted by a standing-wave field, and the stability of the resulting stationary nonequilibrium states of the paramagnet is analyzed. © *1999 American Institute of Physics.* [S1063-7842(99)01805-X]

In recent years the investigation of nonlinear dynamical phenomena in nonequilibrium dissipative laser-type systems has borne a number of fundamental results associated with the detection of bistability, deterministic chaos, optical turbulence, and self-organization in stimulated photon emission.[1–3] At this stage, however, there is a growing awareness of a shortcoming inherent to lasers operating in the optical range: a high spontaneous emission level, which generates essentially irremovable multiplicative noise in the investigated system.[4] This feature hinders or renders altogether impossible the experimental observation of many complex deterministic states and motions, which become blurred, distorted, and even obliterated by internal quantum noise. On the other hand, there is a class of stimulated emission systems, such as masers, which operate in the microwave range — in particular, paramagnetic masers that amplify electromagnetic signals[5,6] and phonon masers, including hypersound amplifiers and oscillators[7,8] — for which the spontaneous noise level is approximately 15 orders of magnitude lower than in lasers (owing to the cubic frequency dependence of the spontaneous emission intensity). Consequently, the concepts of optical-range quantum electronics can be exploited to advantage in setting up experiments with microwave dissipative systems, which can be regarded as deterministic under far more general conditions, for example, in a fine stratification ofphase space or in the vicinity of nonequilibrium phase-transition points.[9]

The possibility of observing nonlinear processes in maser amplifiers was first discussed back in the 1960s,[5,10,11] but the nonlinearity parameter B_N obtained in the papers was equal to the product of the pump transition saturation factor Z_P and the ratio $r = \tau_2/\tau_1$, where τ_2 and τ_1 are the transverse and longitudinal spin relaxation times, respectively. In contrast with laser active media, where r can be of the order of unity, the paramagnets used in microwave-range quantum amplifiers have typical values $r = 10^{-6} - 10^{-8}$. This characteristic hinders actual observation of the effects predicted in Refs. 5, 10, and 11, because it is difficult to attain saturation

at large values of Z_P without overheating the crystal, and B_N must be of the order of unity. On the other hand, experiments on real quantum amplifiers[7,12,13] have shown that in fact appreciable deviations from linearity are already observed for $B_N \ll 1$, indicating the prevalence of a qualitatively different nonlinearity mechanism with no need for such deep saturation of the spin system ($Z_P \simeq 10^6 - 10^8$) as required by the model in Refs. 5, 10, and 11.

A model of resonance saturation of a paramagnet has been proposed,[14,15] taking into account the self-interaction of the saturating field through a system of active centers in a microwave cavity for $q_1 \equiv 2\tau_c/\tau_1 \ll 1$, where τ_c is the photon lifetime in the free microwave cavity. Unlike the older model,[5,10,11] here the nonlinearity parameter is defined as $\xi \equiv q_2 \xi_0 = Q_c^{(0)}/Q_m$, where $q_2 = 2\tau_c/\tau_2$, ξ_0 is the renormalized nonlinearity parameter, which does not depend on q_2, $Q_c^{(0)}$ is the intrinsic loaded Q of the microwave cavity outside magnetic resonance regions, and Q_m is the magnetic Q of the investigated dissipative system (paramagnet + microwave cavity) at the pump frequency without saturation of the spin transitions. In Refs. 14 and 15 it has been found that in the case $\xi > 1$ such clearly pronounced nonlinear phenomena as bistability, self-modulation, instability, etc., can occur in a dissipative system of the kind discussed here even for $Z_P \simeq 10 - 10^2$ as a consequence of strong internal feedback in the microwave cavity. Consequently, the conditions for validity of the nonlinear saturation model[14,15] are fully consistent with the true values of the control parameters in maser amplifiers,[5–13] and this fact, in particular, adapts the model quite well to the analysis of previously observed nonlinear dynamical processes.[7,8,12,13]

The most intriguing consideration, however, is the possibility of the onset of branched dynamical stationary states in the maser pump system, which have been predicted by the new model,[14,15] but have yet to be observed experimentally. Branching of the paramagnetic susceptibility in the pump transition implies that the ordinary monostable regime of stimulated microwave emission[5,6] is transformed into a mul-

tistable regime,[1–3,14,15] which can be accompanied by a hysteresis-type dependence of the maser gain G on certain control parameters, for example, on the pump input power $\widetilde{P}$, the magnetic field detuning of the spin system $\Delta H \equiv H - H_0$ relative to the peak of the pump resonance line H_0, etc. The main objective of the present study is to conduct an experimental search for conditions under which such nonlinear effects take place in a paramagnetic maser.

The investigations were carried out on a millimeter-range maser with an active crystal of andalusite $Al_2SiO_5:Fe^{3+}$ (Ref. 6), in which additional conditions were established for efficient self-interaction of the pump field through a system of spins precessing in a static magnetic field **H**. The field **H** was oriented along the crystallographic **c** axis of the andalusite. Owing to the strong zero-field splitting of the spin doublets of the Fe^{3+} ion in andalusite (116.1 GHz between the first and second doublets and 232.2 GHz between the second and third doublets), at a pump frequency $\Omega_P = 143.6$ GHz the resonance field $H = H_0$ is only about 5.4 kOe, and the frequency of the amplified signal field (used to record the inversion) is, accordingly, $\Omega_S = 43.5$ GHz.

Previous estimates[15] of the control parameters for such a maser have shown that the threshold for the onset of bistability in the pump system can be greatly exceeded at liquid helium temperatures if the system operates in the nonlinear microwave cavity regime. In the signal channel, on the other hand, it is advisable to maintain a traveling-wave regime to prevent ''intrinsic'' nonlinearities in the system recording the gain of the test field Ω_S. To ensure optimal conditions for interaction of the amplified field with the active medium, a finger-array retarding system was employed with a group-velocity retardation factor as high as 50. Stable operation of the maser in the signal channel was provided by a nonreciprocal element made of a W-type hexaferrite.[6]

To establish conditions for efficient self-interaction of the pump field Ω_P while preserving the traveling-wave regime for the recording signal field Ω_S, the polarizations of these fields were set to be mutually perpendicular, and a selective filter was added to the retarding system. The signal power losses at the selective filter were $L_S < 0.2$ dB, whereas the pump field underwent essentially total secondary reflection from the selective filter with appropriate orientation of the latter. As a result, the maser pump system operated as a microwave cavity with a loaded Q factor $Q_c^{(0)} > 200$. The signal and pump channels were decoupled by means of a microwave Mach–Zehnder interferometer. There was no appreciable saturation of the Fe^{3+} spin system by the signal field, so that the experimentally measured gain G was proportional to the inversion ratio K, which qualitatively characterizes the stationary nonequilibrium state of the given dissipative system.

At cryostat temperatures of 2–4.2 K in the pump power range $\widetilde{P} \approx 0.1$–0.35 W we observed branching of the inversion ratio against the background of the total nonmonotonic dependence $K(\widetilde{P})$, which was manifested as hysteresis of the maser gain during slow scanning of the pump power. Characteristically, when the cryostat temperature was lowered
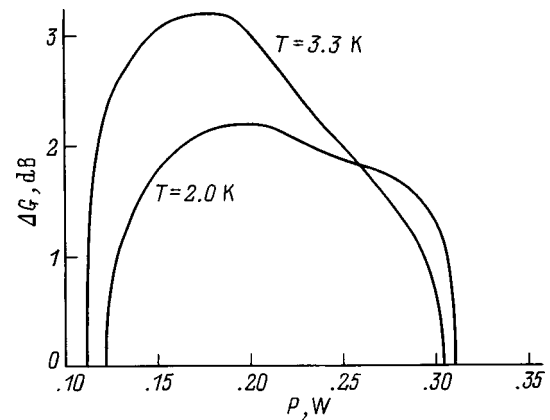

FIG. 1. Experimental plots of the width ΔG of the hysteresis loop as a function of the pump power $\widetilde{P}$.

from 3.3 K to 2 K, not only did the hysteresis loop not broaden, its even became considerably narrower (Fig. 1), clearly revealing the nonthermal origin of the hysteresis. Also, since very low pump power scanning rates were used in our experiments ($d\widetilde{P}/dt < 0.1$ dB/min), neither ordinary (lattice-type) nor spin thermodynamic transition processes contributed to the formation of the $K(\widetilde{P})$ hysteresis loops in these experiments.

The scanning of the pump power was interrupted with test pauses ($d\widetilde{P}/dt = 0$) lasting at least 3 min with G being continuously recorded both for transmission of the upper branch and for transmission of the lower branch of the hysteresis loop (at each of the working temperatures $T = 33$ K and 2 K). The observed gain G remained constant during such test pauses within the error limits of our measurements (± 0.2 dB). The only exception occurred in the very narrow tip parts $\Delta\widetilde{P}$ of the hysteresis loops, where delayed switching of the inversion branches took place, probably on account of the critical slowing down[1] in the given dissipative system as the control parameter $\widetilde{P}^{(\text{start})}$ approached the bifurcation value $\widetilde{P}^{(\text{cr})}$.

In general, the hysteresis processes in devices of the maser type can be attributed both to transient effects in monostable pump regimes, as observed previously in experiments with a ruby ($Al_2O_3:Cr^{3+}$) phonon maser,[7] and also to the formation of bistable (multistable) stationary nonequilibrium states of such systems, which has been justified theoretically.[14,15] Hysteresis in a monostable system always exhibits transient behavior and, of course, occurs only in certain intervals control parameter scanning rates. For example, the hysteresis of the gain of the phonon signal in H scanning as observed in a ruby phonon maser is associated with a very long relaxation time in the electron–nuclear system of $Al_2O_3:Cr^{3+}$ (of the order of 10 s at $T = 1.7$ K). For sufficiently slow scanning of the control parameter the hysteresis loop in a monostable system continuously degenerates, as has been observed[7] for two different configurations for the monostable pumping of Cr^{3+} ions in ruby.

The behavior of a bistable system differs significantly: The more slowly the control parameter is scanned, generally
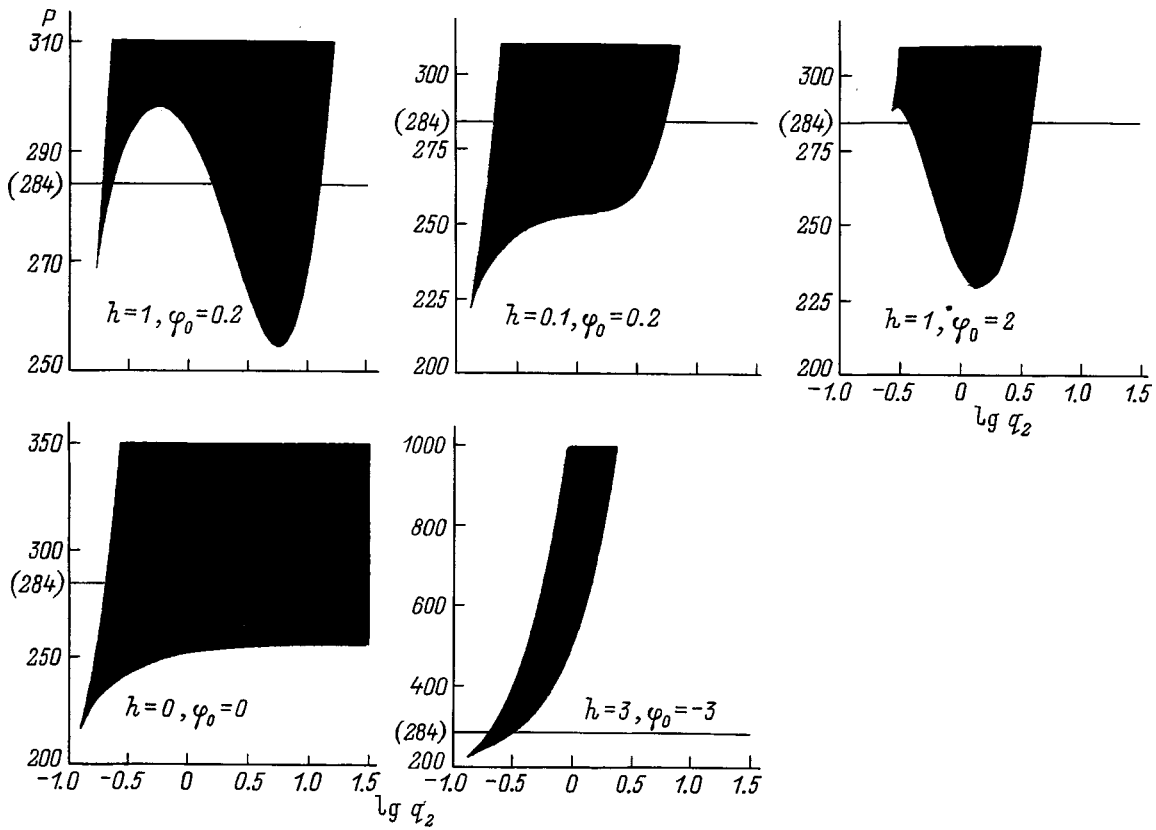
FIG. 2. Typical cross sections of the bifurcation surface in hyperplanes $\{\xi_0, h, \varphi_0\}$ for $\xi_0 = 128$.

speaking, the more favorable are conditions for the onset of hysteresis[1–3] (provided, of course, that the system is structurally stable, and the coexisting branches are not isolated). The only requirements are the absence of strong internal noise and not too pronounced fluctuations of the control parameters, so that the operating point of the system in phase space does not go beyond the limits of the basin of attraction of the attractor corresponding to the scanned branch. Since, as mentioned, spontaneous emission is very weak in amaser, the fluctuations of the control parameters are manageable, and stationary saturated-transition states are recorded from the variation of a nonsaturating test resonance field Ω_S, the results of our experiments described above can be interpreted unequivocally as bistability of stationary, nonequilibrium spin states.

Consequently, the previously proposed[14,15] model of saturation of the spin system in a microwave cavity has been confirmed experimentally, opening the door to a reexamination of the issue of the excitation of stationary inversion states in paramagnetic masers. The most important problem is the optimum ratio of τ_c and τ_2 for the maximization of K on the upper branch of the bistable pump system. To analyze this problem, we refer to expressions describing the surface of saddle–node bifurcations in a maser pump system [Eqs. (5) in Ref. 15]. For convenience we introduce the q_2-independent pump parameter $P \equiv \mathcal{P}/q_2 \propto \tilde{P}$ and plot the cross sections of the bifurcation surface in the hyperplanes $\{\xi_0, h, \varphi_0\}$, where $\varphi_0 \equiv \varphi/q_2$; expressions for the normalized pump power $\mathcal{P}$, the normalized magnetic field detuning of

the spin system h, and the normalized Ω_P-detuning of the microwave cavity frequency are given in Ref. 15. Figure 2 shows the main variants of these cross sections (the filled areas indicate the branching regions of stationary inversion states of the maser in the plane $\{q_2, P\}$, and the heavy horizontal line indicates the identical value of $P = 284$ for all the cross sections). The qualitative difference in these cross sections for $\varphi_0 \neq 0$ and for $\varphi_0 = 0$ are clearly evident. In the first case, as P increases, the region of multiple-valuedness of the function $K(q_2)$ always remains bounded (although not necessarily continuous) in q_2. In contrast, as P increases in the case $\varphi_0 = 0$ (the lower left cross sectionin Fig. 2), the usual q_2-hysteresis part of the function $K(q_2)$ rapidly broadens toward higher values of q_2, changing into a region semibounded in q_2, where we infer from Refs. 14 and 15 that the upper asymptotically stable branch $K_3^{(\varphi_0)}(q_2)$ coexists with the isolated asymptotically stable branch $K_1^{(\varphi_0)}(q_2)$. In this case, since $\partial Z_P / \partial P < 0$, the branch $K_2^{(\varphi_0)}(q_2)$ is always unstable and is a separatrix for the basins of attraction of the attractors $A_{\text{low}}^{(0)}$ and $A_{\text{high}}^{(0)}$ corresponding to the branches $K_1^{(\varphi_0)}(q_2)$ and $K_3^{(\varphi_0)}(q_2)$.

It can be shown that the branch $K_3^{(\varphi_0)}(q_2)$ is a continuation of the single-valued function $K^{(\varphi_0)}(q_2)$ after the operating point intersects the bifurcation surface, where now $\partial K^{(\varphi=0)}/\partial q_2 > 0$ over the entire region of variation of q_2. This result is fully consistent with the intuitive notion that the higher the Q of the pump cavity (i.e., the higher the value of q_2), the larger will be the inversion ratio at a fixed pump
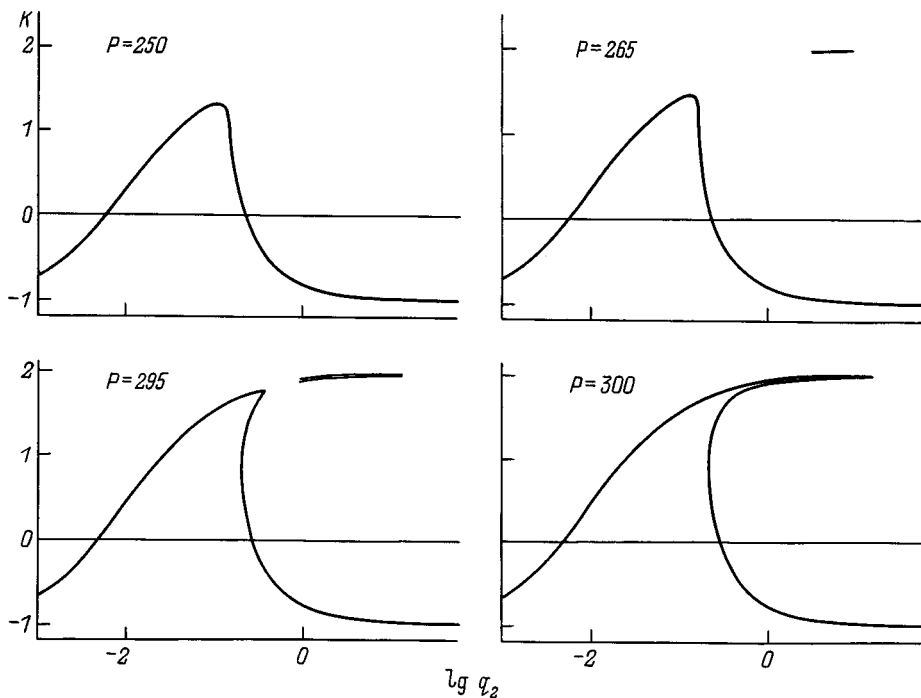
FIG. 3. Nonlinear resonance of the inversion ratio with respect to the parameter q_2 for $\xi_0 = 128$, $\varphi_0 = 0.2$, $(\partial \Omega_P / \partial H)_{H_0}(H - H_0)\tau_2 = 1$ for a spin-system with the maximum $K = K_{\max} = 2$.

power. It might appear that the lower branch (usually not an inversion branch) $K_1^{(\varphi_0)}(q_2)$ could be disregarded by virtue of its isolation. In reality, the situation is far more complicated. For example, with the help of Eq. (6) in Ref. 15 we find that as q_2 increases, the branch $K_3^{(\varphi_0)}(q_2)$ presses indefinitely close to the separatrix $K_2^{(\varphi_0)}(q_2)$. Consequently, small but finite perturbations drive the system operating point out of the basin of attraction of the attractor $A_{\text{high}}^{(0)}$, taking the system from the nonequilibrium state to an equilibrium state (with a negative inversion factor). In other words, the inversion state formed in the paramagnet for large q_2 as a result of the in-phase reradiation of the pump field in the microwave cavity corresponds to a very shallow local minimum of the potential.[1] Conversely, for the same values of q_2 the inversion state corresponding to antiphase reradiation of the pump field is stable even against large perturbations; the pump energy is spent not on inversion but in maintaining this absorption regime.

Consequently, the inclination to keep increasing the Q of the microwave cavity with the idea of lowering the pump power produces the opposite result: collapse of the inversion state due to the inopportune growing proximity of $A_{\text{high}}^{(0)}$ to the separatrix. Moreover, "exact tuning of the cavity" $\varphi_0 = 0$ *per se* is an exceptional situation in that the slightest deviation of the pump source from the natural frequency of the microwave cavity converts the indicated semibounded region of the cross section of the bifurcation surface into a bounded (possibly discontinuous) region in the direction of the q_2 axis. This situation corresponds to structural instability of the inversion states formed as a result of strict in-phase reradiation of the pump field at high values of the microwave cavity Q.

Figure 3 shows graphs of $K(q_2)$ for $\varphi \neq 0$ at various pump levels for the same set of control parameters. The resonance character of these curves is clearly visible: The para-

magnet acquires a stationary nonequilibrium state, which attains maximum values and smoothly or discontinuously decays in a certain interval (or intervals) of q_2. It is evident from Fig. 3 that this nonlinear resonance is characterized by a single-valued or multiple-valued dependence of K on q_2; in a certain interval of q_2 the inversion state resides very close to the separatrix (intermediate branch) and, not only that, also contains isolated subintervals in which forward and backward saddle–node bifurcations correspond to the same pair of branches of the inversion ratio. Consequently, the actual width of the observed resonance in these cases is substantially smaller than the width determined from the end points of the intervals in which $K \geqslant 0$.

In summary, we have experimentally observed multiple-valued, nonlinear dependences of the inversion ratio of the spin system on the pump power, and we have proposed an interpretation of this phenomenon based on the model of resonance saturation of the paramagnet in the presence of self-interaction of the pump field in the microwave cavity,[14] with the prediction of branching of stationary nonequilibrium states of the paramagnetic spin system.[14,15] Based on an analysis of the $K(q_2)$ curves obtained from the model of Refs. 14 and 15, we have demonstrated the feasibility of correctly choosing the parameters of the cavity pump system so as to obtain stable nonequilibrium states having the maximum possible inversion ratio.

[1] H. M. Gibbs, *Optical Bistability: Controlling Light With Light* (Academic Press, Orlando, 1985).

[2] P. Berge, Y. Pomeau, and C. Vidal, *L'Ordre Dans le Chaos: Vers une Approche Determinsite de la Turbulence* (Hermann, Paris, 1988).

[3] G. Haken, *Information and Self-Organization: A Macroscopic Approach to Complex Systems* (Springer-Verlag, Berlin-Heidelberg, 1988).

[4] W. Horsthemke and R. Lefever, *Noise-Induced Transitions: Theory and Applications in Physics, Chemistry, and Biology* (Springer-Verlag, Berlin, 1984).

[5] A. A. Vuylsteke, *Elements of Maser Theory* (Van Nostrand, New York, 1960).

[6] A. A. Lavrinovich, T. S. Smirnova, N. T. Cherpak *et al.*, Dokl. Akad. Nauk SSSR **297**, 857 (1987) [Sov. Phys. Dokl. **32**, 1015 (1987)].

[7] D. N. Makovetskiĭ, Author's Abstract of Candidate's Dissertation [in Russian], Kharkov, 1984.

[8] E. M. Ganapol'skiĭ and D. N. Makovetskiĭ, Pis'ma Zh. Tekh. Fiz. **20**(21), 65 (1994) [Tech. Phys. Lett. **20**(11), 882 (1994)].

[9] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1984).

[10] A. M. Clogston, J. Phys. Chem. Solids **4**, 271 (1958).

[11] A. Jelenski and H. Szymczak, Bull. Acad. Pol. Sci. Ser. Sci. Tech. **15**, 43 (1967).

[12] H. Inaba and H. Morita, in *Physics of Quantum Electronics* (Conference Proceedings), edited by P. L. Kelley (McGraw-Hill, New York, 1966).

[13] N. T. Cherpak, Ukr. Fiz. Zh. **20**, 1715 (1975).

[14] D. N. Makovetskiĭ, Ukr. Fiz. Zh. **33**, 1237 (1988).

[15] D. N. Makovetskiĭ and K. V. Vorsul', Zh. Tekh. Fiz. **61**(1), 86 (1991) [Sov. Phys. Tech. Phys. **36**, 50 (1991)].

Translated by James S. Wood

Motion of a particle in a static magnetic field and the field of a monochromatic electromagnetic plane wave

E. M. Boldyrev

Institute of High-Energy Physics, 142284 Protvino, Moscow Region, Russia
(Submitted January 8, 1998; resubmitted October 19, 1998)
Zh. Tekh. Fiz. **69**, 106–110 (May 1999)

A solution of the equation of motion of a charged particle in an external electromagnetic field comprising a superposition of a uniform static magnetic field and the field of a monochromatic, elliptically polarized electromagnetic plane wave is obtained as the solution of a Cauchy problem. The resonance case is investigated. An analysis of the resulting solution is given. © *1999 American Institute of Physics.* [S1063-7842(99)01905-4]

INTRODUCTION

A solution of the equation of motion of a charged particle in an external field comprising a superposition of a uniform static magnetic field and the field of a monochromatic, circularly polarized electromagnetic plane wave has been obtained previously[1] as the solution of a Cauchy problem. In this paper we find an analogous solution for a monochromatic, elliptically polarized electromagnetic plane wave, encompassing as special cases the above-indicated solution and the solution for a linearly polarized, monochromatic electromagnetic plane wave.

Apart from its purely academic importance, the problem is also of practical significance in that the indicated type of wave very accurately matches the electromagnetic field of a laser beam used as the basis of operation of such systems as the free-electron laser and the undulator.[2] The development of these systems requires a careful analysis of charged-particle motion in the given field configuration.

The actual statement of the problem (the primary electromagnetic wave is represented in its most general form), the correctness of the solution, and the final solution itself expressed as an explicit dependence on the initial data, the polarization amplitudes, and a combination of the sign of the particle charge and the degree of polarization of the electromagnetic wave are such that the solution can be used directly in practical calculations.

CONDITIONAL NOTATION AND STATEMENT OF THE PROBLEM

In this paper (x,y,z,ct) are the coordinates of a point in four-dimensional space. The three-dimensional V vector in this space is denoted, as usual, by $\mathbf{V}$, $[x,y,z]$ are the coordinates of this vector, $\mathbf{a} \cdot \mathbf{b}$ denotes the scalar product of vectors, $\mathbf{r}$ is the radius vector of a charged particle, $\mathbf{r}_0$ is the radius vector of its initial position, m is the mass of the particle, ε is the kinetic energy of the particle, ε_0 is its initial energy, $\mathbf{P}$ is the momentum of the particle, $\mathbf{P}_0$ is its initial momentum, $\mathbf{v}$ is the velocity of the particle, $|e|$ is the charge of the particle, c is the speed of light, φ^W is the scalar potential of the electromagnetic field of the wave, $\mathbf{A}^W$ and $\mathbf{A}^H$ are the vector potentials of the superimposed fields, $\mathbf{A}_0^W$ is

the (complex) amplitude of the electromagnetic field, and $\mathbf{n}$ is the unit vector in the direction of wave propagation. The wave 4-vector for the electromagnetic field of the wave is

$$\left(\frac{\omega}{c}, \mathbf{k}\right), \quad \mathbf{k} = \frac{\omega}{c}\mathbf{n},$$

$\mathbf{E}^W$ and $\mathbf{H}^W$ are the electric and magnetic fields of the wave, and $\mathbf{H}$ is a uniform static magnetic field (H is the magnitude of this field).

The gauge conditions for the electromagnetic field of the wave are $\varphi^W = 0$ and $\operatorname{div}\mathbf{A}^W = 0$,

$$\mathbf{A}^W = \frac{1}{2}[\mathbf{A}_0^W e^{i(\mathbf{kr} - \omega t)} + \text{c.c.}]$$

(c.c denotes complex-conjugate terms).

According to

$$\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{H} = \operatorname{curl}\mathbf{A},$$

$$\mathbf{E}^W = \frac{1}{2}[\mathbf{E}_0^W e^{i(\mathbf{kr} - \omega t)} + \text{c.c.}],$$

where

$$\mathbf{E}_0^W = i\frac{\omega}{c}\mathbf{A}_0^W, \quad \mathbf{H}^W = (\mathbf{n} \times \mathbf{E}^W).$$

Following the standard procedure for taking into account the polarization of an electromagnetic wave,[3] we write the amplitude of the wavefield in the form

$$\mathbf{E}_0^W = \operatorname{Re}\mathbf{E}_0^W + i\operatorname{Im}\mathbf{E}_0^W$$

and define the two real vectors

$$\mathbf{E}_0^{(1)} = \operatorname{Re}\mathbf{E}_0^W \cos\theta + \operatorname{Im}\mathbf{E}_0^W \sin\theta,$$

$$\mathbf{E}_0^{(2)} = \operatorname{Re}\mathbf{E}_0^W \sin\theta - \operatorname{Im}\mathbf{E}_0^W \cos\theta,$$

in such a way that

$$\mathbf{E}_0^{(1)} \cdot \mathbf{E}_0^{(2)} = 0,$$

for which the angle θ must be chosen so that

$$\tan 2\theta = \frac{2(\operatorname{Re}\mathbf{E}_0^W)\cdot(\operatorname{Im}\mathbf{E}_0^W)}{(\operatorname{Re}\mathbf{E}_0^W)^2 - (\operatorname{Im}\mathbf{E}_0^W)^2}.$$

We now choose a coordinate system with the x axis directed along the vector $\mathbf{E}_0^{(1)}$, whereupon the vector $\mathbf{E}_0^{(2)}$ is directed along either the positive or the negative y axis; this condition is taken into account by introducing the vector $g\mathbf{E}_0^{(2)}$, where $g = \pm 1$, so that $\mathbf{E}_0^{(2)}$ is directed along the negative y axis for $g = 1$ and is directed along the positive y axis for $g = -1$ (i.e., g is the degree of wave polarization); the z axis is in the direction of wave propagation (and in the direction of the uniform static magnetic field). We introduce the admissible transformation of coordinates $x' = x$, $y' = y$, $z' = z$, $\xi = t - (z/c)$. In this system (dropping the primes) the components of the electromagnetic field of the primary superposition have the form

$$\mathbf{H} = [-gE_0^{(2)}\sin(\omega\xi - \theta),\ E_0^{(1)}\cos(\omega\xi - \theta), H],$$

$$\mathbf{E} = [E_0^{(1)}\cos(\omega\xi - \theta),\ gE_0^{(2)}\sin(\omega\xi - \theta), 0].$$

To determine the momentum $\mathbf{P}(t)$ of the charged particle in this kind of external field and to subsequently calculate its path $\mathbf{r}(t)$ we have the Cauchy problem[3]

$$\frac{d\mathbf{P}}{dt} = e\mathbf{E} + \frac{e}{c}(\mathbf{v}\times\mathbf{H}), \quad \frac{d\varepsilon}{dt} = e\mathbf{E}\cdot\mathbf{v},$$

$$\mathbf{P}(t_0) = \mathbf{P}_0, \quad \varepsilon(t_0) = \varepsilon_0. \tag{1}$$

The above-introduced coordinate transformation on the set of solutions of problem (1) induces the transformation $x'(\xi) = x(t)$, $y'(\xi) = y(t)$, $z'(\xi) = z(t)$, $\xi = t - z(t)c$, where

$$\frac{d}{dt} = \frac{\varepsilon - cP_z}{\varepsilon}\frac{d}{d\xi}.$$

We set $e = g_e|e|$, where $g_e = +1$ for a positively charged particle, and $g_e = -1$ for a negatively charged particle;

$$\omega_1 = \frac{|e|E_0^{(1)}}{mc}, \quad \omega_2 = \frac{|e|E_0^{(2)}}{mc}, \quad \boldsymbol{\pi} = \frac{\mathbf{P}}{mc}, \quad \boldsymbol{\pi}_0 = \frac{\mathbf{P}_0}{mc},$$

$$\varepsilon = \frac{\varepsilon}{mc^2}, \quad \varepsilon_0 = \frac{\varepsilon_0}{mc^2}, \quad \alpha = \varepsilon - \pi_z, \quad \omega_h = \frac{|e|H}{\alpha mc}.$$

We then have

$$\frac{d}{dt} = \frac{\alpha}{\varepsilon}\frac{d}{d\xi},$$

and problem (1) has the form

$$\frac{d\pi_x}{d\xi} = g_e\omega_1\cos(\omega\xi - \theta) + g_e\omega_h\pi_y,$$

$$\frac{d\pi_y}{d\xi} = g_e\omega_2\sin(\omega\xi - \theta) - g_e\omega_h\pi_x,$$

$$\frac{d\pi_z}{d\xi} = g_e\frac{1}{\alpha}[\pi_x\omega_1\cos(\omega\xi - \theta) + g\pi_y\omega_2\sin(\omega\xi - \theta)],$$

$$\frac{d\varepsilon}{d\xi} = g_e\frac{1}{\alpha}[\pi_x\omega_1\cos(\omega\xi - \theta) + g\pi_y\omega_2\sin(\omega\xi - \theta)],$$

$$\boldsymbol{\pi}(\xi_0) = \boldsymbol{\pi}_0, \quad \varepsilon(\xi_0) = \varepsilon_0. \tag{2}$$

It is evident at once from the third and fourth equations of the system (2) that α is an integral of this system, i.e., an integral of motion, and therefore $\alpha = \varepsilon_0 - \pi_{0x}$. Here $\boldsymbol{\pi}$ and $\mathbf{r}$ are related by the equation

$$\boldsymbol{\pi} = \frac{\alpha}{c}\frac{d\mathbf{r}}{d\xi}. \tag{3}$$

Taking Eq. (3) into account, we can write the left sides of the first and second equations of the system (2) in the form

$$\frac{d}{d\xi}\left[g_e\frac{\omega_1}{\omega}\sin(\omega\xi - \theta) + g_e\frac{\alpha}{c}\omega_h y\right],$$

$$-\frac{d}{d\xi}\left[gg_e\frac{\omega_2}{\omega}\cos(\omega\xi - \theta) + g_e\frac{\alpha}{c}\omega_h x\right],$$

and we have two more integrals of motion

$$\Phi_x = \pi_{0x} - g_e\frac{\omega_1}{\omega}\sin(\omega\xi_0 - \theta) - g_e\frac{\alpha}{c}\omega_h y_0,$$

$$\Phi_y = \pi_{0y} + gg_e\frac{\omega_2}{\omega}\cos(\omega\xi_0 - \theta) + g_e\frac{\alpha}{c}\omega_h x_0.$$

From these integrals of motion, taking Eq. (3) into account, we have a system of equations for the determination of $x(\xi)$ and $y(\xi)$:

$$\frac{dx}{d\xi} = g_e\frac{c}{\alpha}\frac{\omega_1}{\omega}\sin(\omega\xi - \theta) + g_e\frac{\alpha}{c}\omega_h y + \frac{c}{\alpha}\Phi_x,$$

$$\frac{dy}{d\xi} = -gg_e\frac{c}{\alpha}\frac{\omega_2}{\omega}\cos(\omega\xi - \theta) + g_e\frac{\alpha}{c}\omega_h x - \frac{c}{\alpha}\Phi_y. \tag{4}$$

Differentiating the first equation of this system with respect to ξ and taking into account the second equation for the determination of $x(\xi)$, we have the linear homogeneous differential equation

$$\frac{d^2x}{d\xi^2} + \omega_h^2 x = \frac{c}{\alpha}g_e\omega_h\left[\left(\frac{\omega_1}{\omega_h} - gg_e\frac{\omega_2}{\omega}\right)\cos(\omega\xi - \theta) + \Phi_y\right]. \tag{5}$$

In solving this equation, we need to distinguish four cases: $|\omega_h| \neq |\omega|$, $|\omega_h| = |\omega|$, $\omega_h = 0$, $\omega = 0$. The last two cases are treated as special cases of the first. In the first two cases we have the total integral of Eq. (5), $x(\xi, C_1, C_2)$, whereupon we determine $y(\xi, C_1, C_2)$ from (4) and then, using Eq. (3)), find $\pi_x(C_1, C_2)$ and $\pi_y(C_1, C_2)$. We find $\pi_z(\xi, C_1, C_2, C_x')$ (C_x' is additive) and $z(\xi, C_1, C_2, C_z', C_z)$ (here C_z is additive) by elementary integration from the third equation of the system (2). Since only three equations are independent in the Cauchy problem (2) (Ref. 4), from the initial conditions we have six equations to evaluate the four arbitrary constants C_1, C_2, C_z', and C_z, but only four of the six are independent, and we can therefore solve four equations in four unknowns to evaluate the indicated constants and thus to solve the problem.

In the third case the first two equations in (2) are independent, and the system is therefore solved by elementary

integration without using the integrals of motion Φ_x and Φ_y; all six constants of integration are additive and are readily evaluated from the initial conditions of the Cauchy problem (2).

In the fourth case the integration of the equations and the evaluation of the constants are described in detail in the relevant literature (see, e.g, Ref. 4) and will not be given here. It is therefore unnecessary to analyze this solution.

SOLUTION OF THE PROBLEM

The indicated solutions are given below.

Case $|\omega_h| \neq |\omega|$

$x = C_{11} \cos \omega_h(\xi - \xi_0) + C_{12} \sin \omega_h(\xi - \xi_0)$
$\quad + C_{13} \cos (\omega\xi - \theta) + C_{14}$,

$y = C_{21} \cos \omega_h(\xi - \xi_0) + C_{22} \sin \omega_h(\xi - \xi_0)$
$\quad + C_{23} \sin (\omega\xi - \theta) + C_{24}$,

$z = C_{31} \cos [\omega_h(\xi - \xi_0) - (\omega\xi - \theta)] + C_{32} \cos [\omega_h(\xi - \xi_0)$
$\quad + (\omega\xi - \theta)] + C_{33} \sin [(\omega\xi - \xi_0) - (\omega\xi - \theta)]$
$\quad + C_{34} \sin [\omega_h(\xi - \xi_0) + (\omega\xi - \theta)] + C_{35} \sin 2(\omega\xi - \theta)$
$\quad + C_{36}(\xi - \xi_0) + C_{37}$,

$\pi_x = C'_{11} \cos \omega_h(\xi - \xi_0) + C'_{12} \sin \omega_h(\xi - \xi_0)$
$\quad + C'_{13} \sin (\omega\xi - \theta)$,

$\pi_y = C'_{21} \cos \omega_h(\xi - \xi_0) + C'_{22} \sin \omega_h(\xi - \xi_0)$
$\quad + C'_{23} \cos (\omega\xi - \theta)$,

$\pi_z = C'_{31} \cos [\omega_h(\xi - \xi_0) - (\omega\xi - \theta)] + C'_{32} \cos [\omega_h(\xi - \xi_0)$
$\quad + (\omega\xi - \theta)] + C'_{33} \sin [\omega_h(\xi - \xi_0) - (\omega\xi - \theta)]$
$\quad + C'_{34} \sin [\omega_h(\xi - \xi_0) + (\omega\xi - \theta)]$
$\quad + C'_{35} \cos 2(\omega\xi - \theta)] + C'_{36}$.

Case $|\omega_h| = |\omega|$

$x = R_{11}(\xi - \xi_0) \sin (\omega\xi - \theta) + R_{12} \cos \omega(\xi - \xi_0)$
$\quad + R_{13} \sin \omega(\xi - \xi_0) + R_{14}$,

$y = R_{21}(\xi - \xi_0) \cos (\omega\xi - \theta) + R_{22} \cos \omega(\xi - \xi_0)$
$\quad + R_{23} \sin \omega(\xi - \xi_0) + R_{24}$,

$z = R_{31}(\xi - \xi_0)^3 + R_{32}(\xi - \xi_0)^2 + R_{33}(\xi - \xi_0)$
$\quad + R_{34}(\xi - \xi_0) \cos 2(\omega\xi - \theta) + R_{35} \cos 2\omega(\xi - \xi_0)$
$\quad + R_{36} \sin 2\omega(\xi - \xi_0) + R_{37}$,

$\pi_x = R'_{11}(\xi - \xi_0) \cos (\omega\xi - \theta) + R'_{12} \cos \omega(\xi - \xi_0)$
$\quad + R'_{13} \sin \omega(\xi - \xi_0)$,

$\pi_y = R'_{21}(\xi - \xi_0) \sin (\omega\xi - \theta) + R'_{22} \cos \omega(\xi - \xi_0)$
$\quad + R'_{23} \sin \omega(\xi - \xi_0)$,

$\pi_z = R'_{31}(\xi - \xi_0)^2 + R'_{32}(\xi - \xi_0) + R'_{33}(\xi - \xi_0) \sin 2(\omega\xi - \theta)$
$\quad + R'_{34} \cos 2(\omega\xi - \theta) + R'_{35} \sin 2(\omega\xi - \theta) + R'_{36}$.

Case $\omega_h = 0$

$x = W_{11}(\xi - \xi_0) + W_{12} \cos (\omega\xi - \theta) + W_{13}$,

$y = W_{21}(\xi - \xi_0) + W_{22} \cos (\omega\xi - \theta) + W_{23}$,

$z = W_{31}(\xi - \xi_0) + W_{32} \sin 2(\omega\xi - \theta) + W_{33} \cos (\omega\xi - \theta)$
$\quad + W_{34} \sin (\omega\xi - \theta) + W_{35}$,

$\pi_x = W'_{11} \sin (\omega\xi - \theta) + W'_{12}$,

$\pi_y = W'_{21} \cos (\omega\xi - \theta) + W'_{22}$,

$\pi_z = W'_{31} \cos 2(\omega\xi - \theta) + W'_{32} \cos (\omega\xi - \theta)$
$\quad + W'_{33} \sin (\omega\xi - \theta) + W'_{32}$.

Case $\omega = 0$

$x = H_{11} \cos \omega_h(\xi - \xi_0) + H_{12} \sin \omega_h(\xi - \xi_0) + H_{13}$,

$y = H_{21} \cos \omega_h(\xi - \xi_0) + H_{22} \sin \omega_h(\xi - \xi_0) + H_{23}$,

$z = H_{31}(\xi - \xi_0) + H_{32}$,

$\pi_x = H'_{11} \cos \omega_h(\xi - \xi_0) + H'_{12} \sin \omega_h(\xi - \xi_0)$,

$\pi_y = H'_{21} \cos \omega_h(\xi - \xi_0) + H'_{22} \sin \omega_h(\xi - \xi_0)$,

$\pi_z = H'_{31}$.

In every case we have $\varepsilon(\pi_z - \pi_{0z}) + \varepsilon_0$. The coefficients, which we call (not altogether correctly) amplitudes, C_{ij}, C'_{ij}, R_{ij}, R'_{ij}, W_{ij}, W'_{ij}, H_{ij}, and H'_{ij}, are given in Ref. 5. The numerical values of the subscripts ij are evident at once from the given solution.

ANALYSIS OF THE SOLUTION

It is evident from the solution itself and from the amplitude expressions that the motion of a charged particle in the indicated superposition of fields is a fairly complex process and in general is not obtained as the simple superposition of its motions in the electromagnetic field and in the uniform static magnetic field. Consequently, the analysis of the solution (which clearly reduces to an analysis of the amplitudes) is voluminous if one seeks to take into account all the constants on which the amplitudes depend. We shall confine the analysis here to a realm of unquestionable physical interest: the existence of initial and other conditions under which the particle moves in one plane $x = $ const or $y = $ const (without sacrificing generality, we restrict the discussion to the latter plane) and the asymptotic behavior for $\omega \gg 1$.

We say that an expression is zeroth-order if it is $O(1/\omega^0)$ (i.e., does not depend on ω) and that an expression is first-order if it is $O(1/\omega^1)$, etc.

Next, from the explicit expressions for the amplitudes we can regard them conditionally as either independent or such that they can be expressed in terms of the latter. In the ensuing discussion we write the values of the amplitudes in

special cases only for independent amplitudes, with the exception of the asymptotic expressions, because in general this dependence does not contain the orders of the amplitudes.

We note that circular polarization occurs when $\omega_1 = \omega_2$, while in the cases $|\omega_h| \neq |\omega|$ and $|\omega_h| = |\omega|$ the resulting solution is the same as the solution in Ref. 1.

Case $|\omega_h| \neq |\omega|$

We have nine independent amplitudes in this case. For the particle to move in the plane $y = \text{const}$ it is necessary that $\Omega_{\omega\omega_h} = 0$, and it is sufficient to set

$$\pi_{0x} = g_e \frac{\omega_1}{\omega} \sin(\omega\xi_0 - \theta)$$

and $\pi_{0y} = 0$. It follows at once, therefore, that such motion is impossible for circular polarization in the given situation. In this case we have

$$C_{13} = -g_e \frac{c}{\alpha} \frac{\omega_1}{\omega^2}, \qquad C_{14} = x_0 + g_e \frac{c}{\alpha} \frac{\omega_1}{\omega^2} s(\omega\xi_0 - \theta),$$

$$C_{24} = y_0, \qquad C_{35} = -\frac{1}{8} \frac{c}{\alpha^2} \frac{\omega_1^2}{\omega_3},$$

$$C_{36} = \frac{c}{\alpha} \pi_{0z} + \frac{1}{4} \frac{c}{\alpha^2} \frac{\omega_1^2}{\omega^2} \cos 2(\omega\xi_0 - \theta),$$

$$C_{37} = z_0 + \frac{1}{8} \frac{c}{\alpha^2} \frac{\omega_1^2}{\omega_3} \sin 2(\omega\xi_0 - \theta).$$

All other independent amplitudes are equal to zero. For the indicated asymptotic behavior, to within zeroth order, we have

$$C_{11} = H_{11}, \quad C_{12} = H_{12}, \quad C_{14} = H_{13}, \quad C_{21} = H_{21},$$

$$C_{22} = H_{22}, \quad C_{24} = C_{23}, \quad C_{36} = H_{31}, \quad C_{37} = H_{32},$$

$$C'_{11} = H'_{11}, \quad C'_{12} = H'_{12}, \quad C'_{21} = H'_{21}, \quad C'_{22} = H'_{22},$$

$$C'_{36} = H'_{31}.$$

All other amplitudes in this approximation are equal to zero, and we thus have motion in a uniform static magnetic field. To within first-order terms, if $\pi_{0x} = \pi_{0y}$, again we have an oscillatory process with frequency ω_h for the spatial components, but for the moment we have a superposition of two oscillatory processes with frequencies ω and ω_h (cf. Ref. 1).

Note that in the low-frequency limit we have an oscillatory process with frequency ω, which for circular polarization represents helical motion with the indicated frequency.

Case $|\omega_h| = |\omega|$

Here we have 14 independent amplitudes. The nonzero value of R_{11} makes the motion in this case fundamentally different from the motion in the first case. However, the necessary condition for the particle to move in the plane $y = \text{const}$ is in fact $R_{11} = 0$, i.e., the condition $\Omega_- = 0$ or ω_1

$= g g_e \omega_2$, from which it necessarily follows that the case $|\omega_1| \neq |\omega_2|$ be excluded, and we actually have the case of circular polarization. Upon satisfaction of the additional conditions $\pi_{0x} = g_e(\omega_1/\omega)\sin(\omega\xi_0 - \theta)$ and $\pi_{0y} = 0$ we then have

$$R_{12} = -g_e \frac{c}{\alpha} \frac{\omega_1}{\omega^2} \cos(\omega\xi_0 - \theta),$$

$$R_{13} = g_e \frac{c}{\alpha} \frac{\omega_1}{\omega^2} \sin(\omega\xi_0 - \theta),$$

$$R_{33} = \frac{c}{\alpha} \pi_{0z} + \frac{1}{4} \frac{c}{\alpha^2} \frac{\omega_1^2}{\omega^2} \cos 2(\omega\xi_0 - \theta),$$

$$R_{35} = -\frac{1}{8} \frac{c}{\alpha^2} \frac{\omega_1^2}{\omega^2} \sin 2(\omega\xi_0 - \theta),$$

$$R'_{12} = g_e \frac{\omega_1}{\omega} \sin(\omega\xi_0 - \theta), \qquad R'_{13} = g_e \frac{\omega_1}{\omega} \cos(\omega\xi_0 - \theta).$$

All other independent amplitudes are equal to zero.

It is not necessary to say anything about the asymptotic behavior in this case, because ω, which is actually the Larmor frequency, is now commensurate with ω_i ($i = 1, 2$). We can state, however, that to within first-order terms and upon satisfaction of certain conditions, viz., $\Omega_- = 0$, a very high longitudinal momentum, and the vanishing of all other independent amplitudes in this approximation, we have

$$R_{12} = H_{11}, \quad R_{13} = H_{12}, \quad R_{14} = H_{13}, \quad R_{22} = H_{21},$$

$$R_{23} = H_{22}, \quad R_{24} = H_{23}, \quad R_{33} = H_{31}, \quad R_{37} = H_{32},$$

i.e., the paths of the particle are the same as the paths in a uniform static magnetic field. But the same is not true of the momentum, and oscillations of the electromagnetic wave are superimposed on the motion in a uniform static magnetic field.

Case $\omega_h = 0$

We have 11 independent amplitudes in this case. For the particle to move in the plane $y = \text{const}$ it is necessary that $\omega_2 = 0$ and, assuming in addition that $\pi_{0y} = 0$, we see that the amplitudes W_{11}, W_{12}, W_{13}, and W'_{11} do not change, and the remaining nonzero independent amplitudes are equal to

$$W_{23} = y_0,$$

$$W_{31} = \frac{c}{\alpha}\left[\pi_{0z} - g_e \frac{1}{\alpha} \frac{\omega_1}{\omega} \pi_{0x} \sin(\omega\xi_0 - \theta) \right.$$

$$\left. - \frac{1}{4} \frac{1}{\alpha} \frac{\omega_1^2}{\omega^2} \cos 2(\omega\xi_0 - \theta) + \frac{1}{2} \frac{\omega_1^2}{\omega^2} \right],$$

$$W_{32} = -\frac{1}{8} \frac{c}{\alpha^2} \frac{\omega_1^2}{\omega^3},$$

$$W_{35} = z_0 + g_e \frac{c}{\alpha^2} \frac{\omega_1}{\omega^2} \pi_{0x} \cos (\omega\xi_0 - \theta)$$

$$- \frac{3}{8} \frac{c}{\alpha^2} \frac{\omega_1^2}{\omega^3} \sin 2(\omega\xi_0 - \theta).$$

Another interesting consideration in this case is the existence of initial and other conditions such that helical motion occurs, specifically:

$$\pi_{0x} = g_e \frac{\omega_1}{\omega} \sin (\omega\xi_0 - \theta), \quad \pi_{0y} = -g g_e \frac{\omega_2}{\omega} \cos (\omega\xi_0 - \theta).$$

Here the amplitudes W_{12}, W_{13}, W_{22}, W_{23}, W_{32}, W'_{11}, and W'_{21} do not change, and the remaining nonzero amplitudes have the form

$$W_{31} = \frac{c}{\alpha} \left[\pi_{0z} + \frac{1}{4} \frac{1}{\alpha} \frac{\omega_1^2 - \omega_2^2}{\omega^2} \cos 2(\omega\xi_0 - \theta) \right],$$

$$W_{35} = z_0 + \frac{1}{8} \frac{c}{\alpha^2} \frac{\omega_1^2 - \omega_2^2}{\omega^2} \sin 2(\omega\xi_0 - \theta),$$

i.e., we have pure helical motion only in the case of circular polarization.

The asymptotic behavior in the zeroth approximation is extremely simple: uniform rectilinear motion. To within first-order terms the amplitudes W_{11}, W_{21}, W'_{11}, W'_{12}, W'_{21}, and W'_{22} do not change, and the remaining amplitudes with nonzero values in this approximation have the form

$$W_{13} = x_0, \quad W_{23} = y_0,$$

$$W_{31} = \frac{c}{\alpha} \left[\pi_{0z} - g_e \frac{1}{\alpha} \frac{\omega_1}{\omega} \pi_{0x} \sin (\omega\xi_0 - \theta) \right.$$

$$\left. + g g_e \frac{1}{\alpha} \frac{\omega_2}{\omega} \pi_{0y} \cos (\omega\xi_0 - \theta) \right],$$

$$W_{35} = z_0, \quad W'_{32} = -g g_e \frac{1}{\alpha} \frac{\omega_2}{\omega} \pi_{0y},$$

$$W'_{33} = g_e \frac{1}{\alpha} \frac{\omega_1}{\omega} \pi_{0x},$$

$$W'_{34} = \pi_{0z} - g_e \frac{1}{\alpha} \frac{\omega_1}{\omega} \pi_{0x} \sin (\omega\xi_0 - \theta)$$

$$+ g g_e \frac{1}{\alpha} \frac{\omega_2}{\omega} \pi_{0y} \cos (\omega\xi_0 - \theta).$$

CONCLUSION

The quantities $\Omega_{\omega\omega_h}$ and Ω_- are defined in Ref. 5. We have obtained a solution of the problem in the form of dimensionless complexes π and ε, but the transition to momentum and energy is perfectly obvious from their definitions.

To transform back from the variables $(x,y,z,c\xi)$ to the variables (x,y,z,ct), it is necessary to solve the nonlinear equation

$$\xi + \frac{z(\xi)}{c} = t.$$

Considerations set forth in Ref. 4 show that a solution $\xi = \xi(t,z)$ of this equation exists and is readily determined numerically. It is simple to find th is solution in the given asymptotic representation, since the particle in fact moves uniformly along the z axis in this case.

In closing, the author wishes to thank A. P. Vorob'ev for a discussion and for comments.

[1] E. M. Boldyrev, Zh. Tekh. Fiz. **67**(2), 94 (1997) [Tech. Phys. **42**, 206 (1997)].

[2] G. Dattoli and A. Torre, in *Free-Electron Laser Theory*, CERN 89-03 (1989).

[3] V. G. Levich, *Course in Theoretical Physics* [in Russian], Vol. 1 (Moscow, 1968).

[4] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 4th ed. [Pergamon Press, Oxford-New York, 1975; Nauka, Moscow, 1973].

[5] E. M. Boldyrev, IFVÉ Preprint No. 97-87 [in Russian], Institute of High-Energy Physics, Protvino, 1997.

Translated by James S. Wood

Dynamics of an electron beam from a high-current picosecond accelerator

K. A. Zheltov, I. G. Zdanovich, and M. N. Nechaev

Scientific-Research Institute of Pulse Technology, 115304 Moscow, Russia
(Submitted January 27, 1998)
Zh. Tekh. Fiz. **69**, 111–115 (May 1999)

The dynamics of a high-current ($10^2 - 10^4$ A) electron beam with energies of $10^5 - 10^6$ eV and picosecond duration (10^{-10} s) at the output of the accelerator tube is investigated. The slowing of electrons by the residual positive charge on the surface of the tube is found to have a significant influence in the case of short pulse durations. The distance of the electron beam from the surface of the tube in vacuum is estimated on the basis of a one-dimensional model. It is shown that the electron radiation can travel to a distance of several centimeters from the surface at current densities below 20 A/mm^2, whereas at high current densities the beam is trapped near the surface. © *1999 American Institute of Physics.* [S1063-7842(99)02005-X]

The electron radiation field of a high-current picosecond accelerator[1,2] is formed in a vacuum diode (tube) and exits through a window transparent to the accelerated electrons. The characteristics of the radiation are determined by the load imposed by the cathode–anode unit of the tube on the coaxial long line supplying voltage to the tube. Owing to the long line, which generally has high resistance ($\sim 50\,\Omega$), the impedance of the cathode–anode gap of the tube is dominated by the resistive component. Evidence of this condition is found in the identical profiles of the incident and reflected pulses in the coaxial line. The resistance of the tube depends, on the one hand, on the area of the emitting surface or the number of emitting centers and, on the other, on the distance between the cathode and the anode.

In addition to the resistive component, the impedance of the tube can also contain a small capacitive component, particularly in the case of the IMAZ-150É tube,[3] where a gradient shield of relatively large size is used to protect the seal between the cathode holder and the glass against breakdown.

The electron energy and the beam current can vary over wide ranges, depending on the ratio between the impedance of the tube and the wave impedance of the line. For high tube impedances the accelerating voltage can attain extremely high levels, tending to a value of twice the amplitude of the pulse traveling in the line. The beam current is not very high in this case. Conversely, a low impedance of the cathode–anode gap corresponds to low electron energies and high currents, ranging up to tens of kiloamperes.

When the tube emits a picosecond (10^{-10} s) electron beam into free space, a special problem encountered is the slowing of electrons by the field of the positive charge induced on the outer surface of the tube. Depending on the magnitude of this retarding field, the electron radiation can either be pulled, extending out to several centimeters, or be squeezed toward the anode, in which case the beam only reaches a few millimeters or even a fraction of a millimeter from the tube.

It is important to note that the electron slowing effect is not discussed anywhere in the literature on the transport of electron beams,[4–7] because the duration of the electron beams is usually greater than 10^{-9} s. For such pulse durations in air or in other media into which the beam is emitted, there is sufficient time for a conductivity to develop, attenuating the retarding field. At pulse durations of 10^{-10} s or less the conductivity does not have time to suppress the retarding field, and the slowing effect is appreciable. In this paper we investigate the electron slowing effect for beam output into vacuum, when suppression of the retarding field by conductivity does not have to be taken into account. At a later time we propose to investigate the output of an electron beam into air with allowance for the conductivity induced in it.

The general relations for the slowing of an electron beam by a positive charge induced on the anode of the tube are best understood in the simplest model of instantaneous emission of a plane layer of electrons with identical velocity from an infinite plane. It can be shown[8,9] that the beam-squeezing Lorentz force in an axisymmetric, relativistic electron beam of finite transverse dimensions is equal in order of magnitude to the repulsive force of the space charge (in the approximation $v \approx c$). We can therefore use the approximation of one-dimensional motion of electrons perpendicular to the plane. Moreover, when the electron layer moves to a distance of the order of the beam diameter, the electric field between the electron layer and the in-plane positive charge can be assumed equal to the field in a plane capacitor:

$$E = q/\varepsilon_0 S, \tag{1}$$

where q is the charge of the emitted electrons, $S = \pi r^2$ is the cross section of the electron beam, r is the beam radius, and $\varepsilon_0 = 8.854 \times 10^{-12}$ F/m is the permittivity of free space.[10]

The retarding electric field can attain huge values for high-current tubes. For example, if the tube current is $I = 1$ kA and the pulse duration is $\tau = 100$ ps, the emitted charge is $q = 1 \times 10^{-7}$ C. For a beam of radius $r = 4$ mm the field is $E = 2.25 \times 10^8$ V/m.

The electron layer is effectively subjected to one-half the field (1).[9] The maximum distance of the electrons from the surface of the tube can be determined from the energy conservation law without solving the equation of motion of relativistic electrons in a static field,

$$L = 2W/e \cdot E = 2U/E, \qquad (2)$$

where $W = e \cdot U$ is the initial kinetic energy of the electrons, U is the accelerating voltage in the tube, and $e = 1.602 \times 10^{-19}$ C is the charge of an electron;[10] at $U = 1$ MV the distance of the electrons from the surface of the tube is ~ 9 mm.

The stopping time of the moving electrons can be determined analogously from the momentum conservation law,

$$\tau = \frac{2p}{eE}, \qquad (3)$$

where p—the initial electron momentum depends on the accelerating voltage of the tube:

$$p = mc \sqrt{\left(1 + \frac{eU}{mc^2}\right)^2 - 1}, \qquad (4)$$

$m = 9.109 \times 10^{-31}$ kg is the electron mass, and $c = 2.9982 \times 10^8$ m/s is the speed of light;[10] at $U = 1$ MV we obtain $\tau = 4.2 \times 10^{-11}$ s.

It follows from Eqs. (1)–(3) that the slowing effect is more pronounced the higher the surface density of the emitted charge q/S or, at a fixed pulse duration, the higher the beam current density $j = I/S$. In the given example we have $j \approx 20$ A/mm^2.

A shortcoming of the single-layer model is the assumption that electrons are instantaneously emitted with identical velocities, so that the layer subsequently moves as a unit whole without breaking apart. In reality, the electron pulse has a finite duration, and the accelerating voltage, which dictates the initial electron velocity, varies during the life of the pulse.

In the multilayer model of electron beam slowing the bell-shaped time dependence of the current is approximated by the equation

$$I(t) = \frac{1}{2} I_{\max} \left[1 - \cos\left(\frac{\pi}{T} t\right) \right], \quad 0 \leqslant t \leqslant 2T, \qquad (5)$$

where $I_{\max}$ is the amplitude of the current, and T is the pulse duration at half maximum (FWHM); the duration of the pulse at its base is $2T$.

The accelerating voltage is approximated by the analogous equation

$$U(t) = \frac{1}{2} U_{\max} \left[1 - \cos\left(\frac{\pi}{T} t\right) \right], \quad 0 \leqslant t \leqslant 2T, \qquad (6)$$

where $U_{\max}$ is the amplitude of the accelerating voltage.

The motion of the beam beyond the limits of the tube is calculated as follows. The base-line pulse duration $2T$ is partitioned into a number of identical time subintervals of width Δt_i. Usually 10–20 subintervals is sufficient. The electrons emitted in each time subinterval in the calculations are consolidated in a single plane layer. Each layer is assumed to be emitted instantaneously at the midpoint of the subinterval with a velocity corresponding to the accelerating voltage at that time. The electric charge of the layer is

$$q_i = \int_{\Delta t_i} I(t) \cdot dt. \qquad (7)$$

It is readily verified that the total charge Q contained in the current pulse is

$$Q = \int_0^{2T} I(t) \cdot dt = I_{\max} T. \qquad (8)$$

For each layer there is a corresponding field E_i produced by the charge q_i according to Eq. (1).

The procedure used to calculate the effective field $E_{i\,\text{eff}}$ acting on each layer is more complex than in the case of a single layer, especially when allowance is made for the fact that layers having different velocities overtake each other during their motion. Each layer, situated at a distance X_i from the surface of the tube, is subjected to the total field E_j from more-distant layers at $X_j > X_i$, but is not affected by the fields from less-distant layers at $X_j < X_i$, and is subjected to half the field from the given layer i and from layers situated at the same distance as this layer $X_j = X_i$,

$$E_{i\,\text{eff}} = \sum_{X_j > X_i} E_j + \frac{1}{2} \sum_{X_j = X_i} E_j. \qquad (9)$$

In each computational time step the position of each layer X_i at the end of the step and the momentum of the electrons are determined from the integrals of motion of the electrons in the static field $E_{i\,\text{eff}}$. The field $E_{i\,\text{eff}}$ is determined again for a new layer position X_i, and the procedure of calculating $E_{i\,\text{eff}}$, the layer position X_i, and the momenta is repeated cyclically.

Figures 1a–c show the results of calculations of the distances of the electron layers from the surface of the tube for an electron beam of radius $r = 4$ mm and various tube currents $I_{\max}$. The pulse is assumed to be partitioned into 12 subintervals, which is sufficient for obtaining correct computational results. The test of sufficiency of the number of subintervals is considered to be the minimum number above which no change is observed in the behavior of the $X_i(t)$ curves, the plane of the figure merely becoming more densely filled with the $X_i(t)$ curves.

It is evident from Figs. 1a–c that the layers do in fact overtake each other in the course of their motion. Layers with the highest initial electron energy, being emitted close to $t \approx T$ and carrying the maximum electric charge, determine the characteristic distance of the beam from the surface of the tube. Even at the instant of emission these layers are subjected to a substantial retarding field from the preceding electron layers. On the other hand, the layers of lower energy emitted from the tube during an earlier stage of the process, when the retarding field is not as strong, move to much greater distances, as, for example, in the case of the third layer ($t \approx 50$ ps) in all three figures.

In the multilayer electron beam model the outer boundary of electron distances from the surface of the tube is diffuse: Different layers travel to unequal distances. However, the highest-energy layers, which carry the maximum charge, traverse approximately identical (within $\sim 20\%$ scatter) distances. The analytical expression derived for the distance L
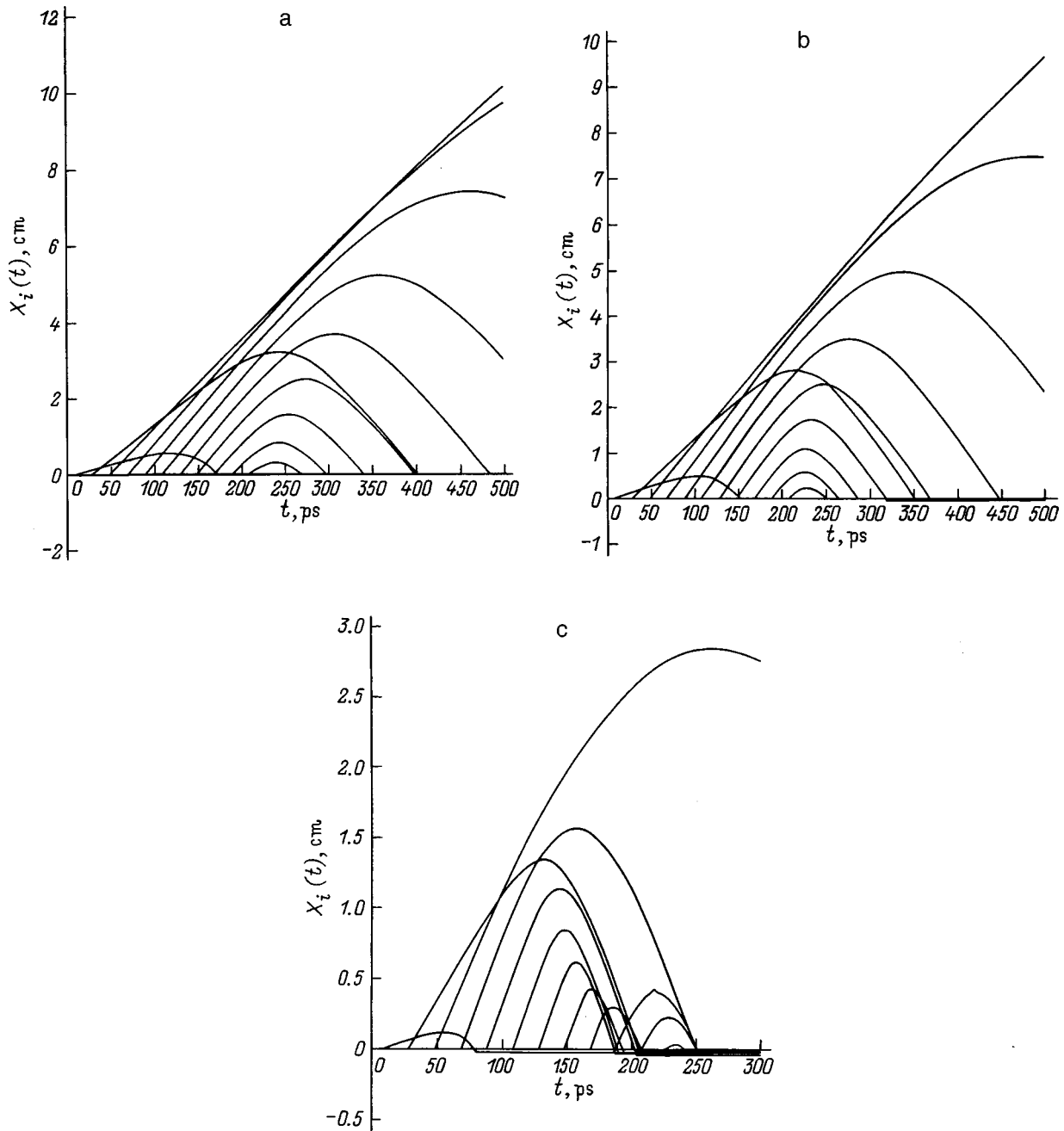
FIG. 1. Distances of layers of the electron beam versus time for an accelerating voltage amplitude $U_{max}=1$ MV, a FWHM pulse duration $T=120$ ps, a beam radius $r=4$ mm, and various characteristic layer distances ($t \approx 0.8T$) and current amplitudes: (a) $L \approx 6$ cm, $I_{max}=0.17$ kA; (b) 4 cm, 0.25 kA; (c) 1 cm, 1 kA.

of the beam from the surface of the tube on the basis of the single-layer model is valid for these characteristic layers if the charge of the layer $q=I_{max}\Delta t$ is assumed to be concentrated in the time interval $\Delta t \approx 2\pi \cdot r/c$, during which light traverses a path equal to the circumference of the beam:

$$L \approx \varepsilon_0 c r \frac{U_{max}}{I_{max}}. \qquad (10)$$

As in the calculations based on the multilayer model, Eq. (10) indicates the occurrence of beam trapping in the tube at sufficiently high currents. For example, for a voltage $U_{max}=1$ MV, a beam radius $r=4$ mm, and beam currents $I_{max}=1$ kA ($j \approx 20$ A/mm^2) the range L is ~ 1 cm, while for a

current amplitude $I_{max}=10$ kA ($j \approx 200$ A/mm^2) the range is $L \sim 1$ mm, i.e., the beam is trapped near the surface of the tube.

Calculations of the energy D transferred by the electron beam to an irradiated surface parallel to the anode surface of the tube at a distance h from it are equally informative here. We consider two limiting cases. In the first case we assume that the electrons incident on the surface no longer impede electrons following them. Physically this assumption implies that the electrons in effect slip off of the irradiated surface onto the tube, i.e., electrical contact is established between the surface and the tube. The results of the calculations for this case are shown in Fig. 2a. The distance from the surface
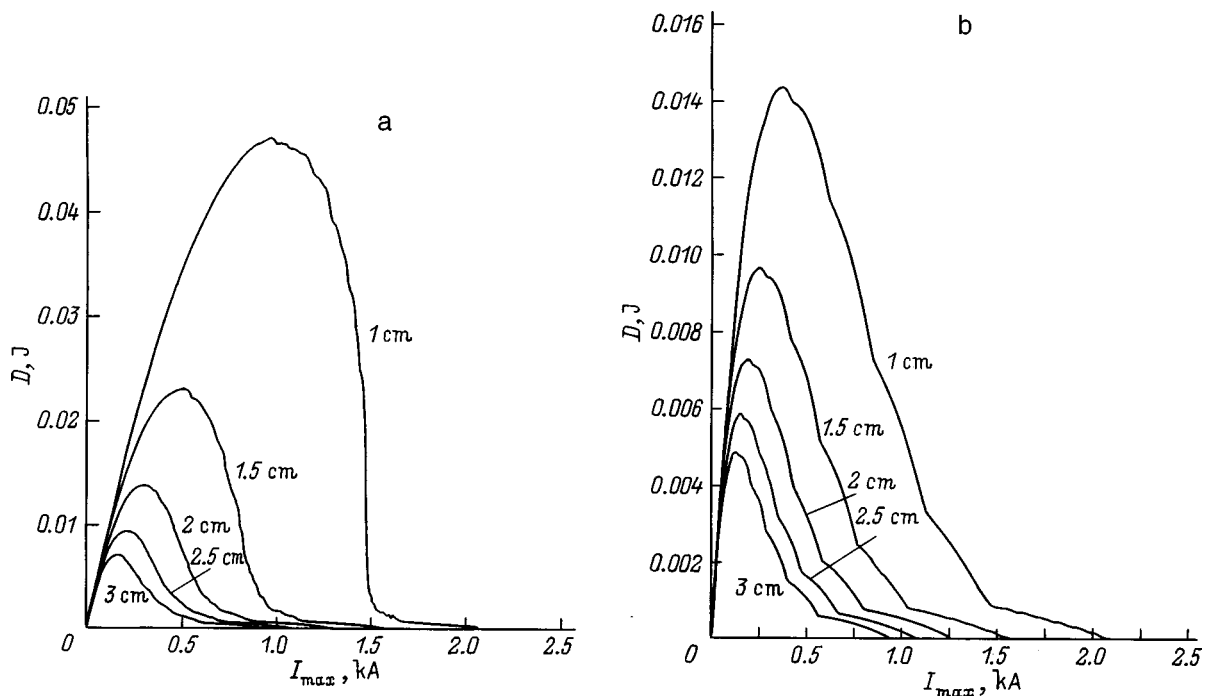
FIG. 2. Energy D transferred by the electron beam to an irradiated surface situated at various distances h from the surface of the tube versus the current amplitude I_{max} for a nonisolated (a) and an isolated (b) irradiated surface.

is varied from 1 cm to 3 cm. In the second limiting case we assume that the electrons incident on the surface stay there and continue to slow down later electrons. Physically this situation corresponds to a dielectric plate isolated from the tube. The results of the calculations are shown in Fig. 2b.

Our estimates of the distance of electron radiation with energy ~ 1 MeV indicate the significant influence of the regarding field generated by the positive charge induced on the outer surface of the anode window. If the electron current of the tube is higher than 1 kA ($j = 20$ A/mm^2), the influence is so strong as to trap the beam close to the surface of the tube. The electron radiation does not have a very long range beyond the limits of the tube in this case, and its energy at a distance of a few centimeters is low.

In the interval of beam current amplitudes of a few hundred amperes (150–300 A) the influence of the retarding field is substantially reduced, and the electron radiation has a range of several centimeters. This range of currents can be regarded as optimal for applications involving the irradiation of external objects.

At low current amplitudes, no more than one or two amperes, the influence of the retarding field is negligible, and the electron radiation has a range of several meters.

In closing, we call attention to a limitation of the given model of electron motion. The basic equation (1) is valid when the electron layer is situated at distances that do not exceed the transverse dimensions of the beam. A major con-

sideration for our discussion is the validity of the model in the investigation of electron beam trapping near the surface of the tube.

[1] G. A. Zheltov, *Picosecond High-Current Electron Accelerators* [in Russian], Énergoatomizdat, Moscow, 1991.
[2] K. A. Zheltov, S. A. Korobkov, A. N. Petrenko, and V. F. Shalimanov, Pribory Tekh. Éksp., No. 1, 37 (1990).
[3] N. G. Pavlovskaya, T. V. Kudryavtseva, S. I. Dron' *et al.*, Pribory Tekh. Éksp., No. 3, 22 (1993).
[4] A. N. Didenko, V. P. Grigor'ev, and Yu. P. Usov, *High-Power Electron Beams and Their Applications* [in Russian], Atomizdat, Moscow, 1977.
[5] A. A. Rukhadze, L. A. Bogdankevich, S. U. Rosinskiĭ, and V. G. Rukhlin, *Physics of High-Current Relativistic Electron Beams* [in Russian], Atomizdat, Moscow, 1980.
[6] M. V. Kuzelev and A. A. Rukhadze, *Electrodynamics of Dense Electron Beams in a Plasma* [in Russian], Nauka, Moscow, 1990.
[7] I. N. Meshkov, *Transport of Charged-Particle Beams* [in Russian], Nauka, Moscow, 1991.
[8] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 4th ed. (Pergamon Press, Oxford-New York, 1975).
[9] I. E. Tamm, *Fundamentals of the Theory of Electricity* [in Russian], Nauka, Moscow, 1966.
[10] *Handbook of Physical Quantities* [in Russian], edited by S. I. Grigor'ev and E. I. Meĭlikhov, Énergoatomizdat, Moscow, 1991.

Translated by James S. Wood

Influence of the initial state of adatoms on the ratio of atoms and ions in a sputtered particle flow

S. Yu. Davydov

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted December 22, 1997)
Zh. Tekh. Fiz. **69**, 116–120 (May 1999)

The role of the initial state of adatoms in the formation of the charge composition of a sputtered
particle flow is investigated. Calculations are carried out for alkali metal atoms adsorbed
on a tungsten surface and subjected to mutual dipole–dipole repulsion. It is shown that the escape
probability (yield) of positive ions diminishes as the surface density of adatoms increases,
an effect that is enhanced when the initial state ofthe adatoms is taken into account. © *1999
American Institute of Physics.* [S1063-7842(99)02105-4]

1. A basic question arises in the theory of electron exchange between a sputtered particle and a metal surface:[1–3] When a particle leaves the surface of the solid in a certain charge state, i.e., as an atom or an ion, does it retain that state or does charge transfer takes place, changing an atom into an ion or changing an ion into an atom? Thequestion of the excitation cross section of the atom or ion is ignored here, and it is immaterial from this standpoint what kind of activity (ion, electron, or neutron bombardment) is responsible for sputtering. The probability of surface charge transfer is described by the occupation number for aparticle escaping to infinity $\langle n(\infty)\rangle$, which has the following significance in the given context: If $\langle n(\infty)\rangle = 1$,the sputtered particle flow consists entirely of atoms, for $\langle n(\infty)\rangle = 0$ we have a flow of singly charged positive ions, and for $0 < \langle n(\infty)\rangle < 1$ the flow contains both atoms and ions. We assume here that only one single-electron orbital is active in charge transfer; this assumption is admissible when the flow contains only atoms (with zero charge) and ions with charge $+1$. It is customarily assumed that the initial state of the particle can be disregarded altogether.[1–3] The reasoning is as follows. In the expression for the occupation number $\langle n(\infty)\rangle$ of a sputtered particle escaping to infinity:

$$\langle n(\infty)\rangle = \langle n(0)\rangle \exp\left[-2\int_{t_0}^{\infty} dt' \cdot \Delta(t') \right]$$
$$+ \pi^{-1} \int d\omega \cdot f(\omega,T) \left| \int_{t_0}^{\infty} dt' \cdot (\Delta(t'))^{1/2} \right.$$
$$\left. \times \exp\left\{ -i\varepsilon t' - \int_{t'}^{\infty} dt''[i\varepsilon^*(t'') + \Delta(t'')] \right\} \right|^2$$

$$(1)$$

($\langle n(0)\rangle$ is the occupation number of the particle at the time of sputtering t_0; 2Δ is the quasi-level width resulting from the finiteness of the electron lifetime on the sputtered particle; $\varepsilon^* = \varepsilon + \Lambda$, where ε is the position of the level of the particle escaping to infinity, and Λ is the quasi-level shift due to interaction with the metal; f is the Fermi distribution function; ω is the energy variable; and the system of units

used here is such that $\hbar = 1$), the first term, which is the so-called memory term in that it specifically contains information on the initial occupation number of the particle, is omitted. Indeed, if the particle is assumed to interact with the substrate during a short time interval τ and if, at the same time, the half-width o f the quasi-level remains approximately constant and equal to Δ_0, the first term on the right-hand side of Eq. (1) has the form

$$\langle n(0)\rangle \exp(-2\Delta_0\tau). \tag{2}$$

Setting $\tau = r_i/v$, where r_i is the ionic radius and v is the normal (to the surface) component of the particle velocity, we find that for $\Delta_0 = 1$ eV and $v = 10^6$ cm/s the first term is of the order of 10^{-14} and can be omitted. In the opposite limit, when the particle is assumed to interact with the surface throughout the entire process, and the width of the quasi-level decreases exponentially with increasing distance from the surface, i.e.,

$$\Delta(z) = \Delta_0 \exp(-2\gamma z) \tag{3}$$

[z is the distance of the particle from the surface, $\gamma \cong 1$ Å^{-1} is a characteristic reciprocal length, and $z = v(t - t_0)$, $v = $ const], the first term has the form

$$\langle n(0)\rangle \exp(-\Delta_0/\gamma v) \tag{4}$$

and is of the order of 10^{-7}. Consequently, the first term of Eq. (1) can be disregarded at velocities of the order of 10^6 cm/s.

However, information on the initial (adsorbed) state of the particle is contained not only in the first term of Eq. (1), but also in the second term, specifically in the quasi-level shift function Λ. As a rule, Λ is interpreted as a hybridized shift, which is related to the broadening Δ by dispersion relations. However, Λ can also be interpreted as the shift of the quasi-level due to other interactions, for example, the interaction of an adatom electron with its image in the metal[4–6] or the dipole–dipole repulsion of adatoms. The latter phenomenon is the subject of this paper.

2. It has been shown[7] that when ions adsorbed on a metal surface are subjected to dipole–dipole repulsion, the effect of

TABLE I. Initial data and parameters of the model.

Parameter	Li	Na	K	Rb	Cs
l, Å	1.13	1.39	1.85	1.99	2.24
d, Å	3.02	3.66	4.53	4.84	5.24
ξ, eV	12.02	10.21	9.57	9.05	9.07
$N_{ML}^{-1} \times 10^{15}$ cm^2	0.91	1.34	2.05	2.34	2.74
Ω_m, eV	-0.14	0.11	0.91	1.07	1.36
Δ_0, eV	2.16	1.47	0.96	0.84	0.72
Φ, eV	22.4	18.8	16.3	15.4	14.8

the latter on an individual ion can be described by an effective field that imparts a shift Λ_0 to the adatom quasi-level:

$$\Lambda_0 = \xi \cdot \Theta^{3/2}(1-n), \quad \xi = 2e^2 l^2 N_{ML}^{3/2} A, \tag{5}$$

where $2l$ is the arm of the dipole formed by the charge of the ion and its image in the metal, N_{ML} is the density of adatoms in the monolayer (subscript ML), $\Theta = N/N_{ML}$ is the relative density of adatoms (coverage), e is the positron charge, and $A \cong 9$ is a coefficient that depends weakly on the geometry of the adsorbed layer.

We note that the approach developed in Ref. 7 is valid, generally speaking, only up to coverages corresponding to the minimum work function of the adsorbed system, because it does not take metallization into account (see below).

The occupation number of an adsorbed adatom $n \equiv \langle n(0) \rangle$ is determined self-consistently on the basis of the Anderson–Newns Hamiltonian from the equation

$$n = \frac{1}{\pi} \arctan \frac{\Omega + \Lambda_0}{\Delta_0}, \tag{6}$$

where $\Omega = \phi - I$ (ϕ is the work function of the substrate, and I is the ionization potential of an adsorbed atom); it is assumed here that not more than one electron can exist at an adatom, i.e., the formation of negative ions is disregarded.

If Eq. (1) is to be used to determine the statistical weights (or escape probabilities) of neutrals (atoms) $P^0 = \langle n(\infty) \rangle$ and positive ions $P^+ = 1 - \langle n(\infty) \rangle$ in the sputtered particle flow, it is necessary to specify the dependence of the quasi-level shift Λ of an escaping particle on the time or, equivalently, on the distance. As is customary,[4–6] we set

$$\Lambda = \Lambda_0 \exp(-\lambda z), \tag{7}$$

where λ is a characteristic reciprocal decay length of the electric potential of the surface lattice of dipoles in the direction perpendicular to the surface.

3. To analyze the role of the initial state of an adatom in the formation of the charge distribution of sputtered particles, we consider a specific system: alkali metal atoms adsorbed on the (110) surface of tungsten. The initial data and parameters describing this system are summarized in Table I. The parameter l (half the dipole arm) in Eq. (5) is set equal to half the sum of the ionic and atomic radii of the alkali metal, whose values are given in Ref. 8 (along with the ionization potentials of alkali metal atoms). This definition of l enables us in the first approximation to take into account the transition of the almost ionic state M^+ to the almost atomic state M^0 as the coverage Θ increases. It is also assumed that

$N_{ML} = d^{-2}$, where d is set equal to the distance between nearest neighbors in the bulk of the alkali metal. The values of d are given in Ref. 9. The work function of the tungsten (110) surface is taken from Ref. 8 and is equal to 5.25 eV. To estimate the quasi-level shift, we use an equation derived in Ref. 10:

$$\Delta_0 = \frac{5\pi}{z_s} V_a, \tag{8}$$

where z_s is the number of nearest neighbors in the bulk of the substrate ($z_s = 8$ for tungsten), and V_a is the matrix element of interaction of nearest neighbors in the bulk of the alkali metal; we interpret V_a as Harrison's universal matrix element $V_{ss\sigma} = -1.32(\hbar^2/md^2)$, where m is the electron mass.[11]

4. We now determine the charge composition of the sputtered particle flow. The following consideration must be addressed in this regard. In calculating the occupation number of an atom in the adsorbed state, naturally, the work function must be interpreted as the work function of the metal substrate, i.e., in our case the work function for the W (110) surface. But then when a particle leaves the surface, it immediately ''senses'' the work function of the entire adsorbed system, i.e., the work function of the substrate coated with a film of the alkali metal. This is the work function that must be taken into account in calculating $\langle n(\infty) \rangle$ (Refs. 10 and 12–14). It has been shown[7] that when alkali metals are adsorbed on a metal substrate, the adsorption-induced change in the work function of the substrate $\Delta\phi$ is

$$\Delta\phi = -\Phi \cdot \Theta \cdot (1-n), \quad \Phi = 4\pi e^2 l N_{ML}. \tag{9}$$

The values of Φ are given in Table I. Assuming that the parameter Λ in Eq. (7) is equal to 2γ in Eq. (3) and passing to the low-velocity limit by a procedure similar to that in Ref. 1, we obtain the following equations for $\langle n \rangle \equiv \langle n(\infty) \rangle$:

a) $\bar\Omega \equiv \Omega + \Delta\phi > 0, \quad \langle n \rangle = \dfrac{c}{b} \exp\left(-\dfrac{\Omega}{\varepsilon_0} b\right),$ (10)

b) $\bar\Omega \equiv \Omega + \Delta\phi < 0, \quad \langle n \rangle = 1 - \dfrac{c}{b^*} \exp\left(\dfrac{\bar\Omega}{\varepsilon_0} b^*\right).$ (11)

Here

$$c = \Delta_0/(\Delta_0^2 + \Lambda_0^2)^{1/2}, \quad b = \arctan(\Delta_0/|\Lambda_0|),$$
$$b^* = \pi - b, \quad \varepsilon_0 = \gamma v. \tag{12}$$

Equations (10) and (11) have been derived on the assumption that $\lambda = 2\gamma$, which corresponds to an identical dependence of both the adatom quasi-level width and the dipole–dipole potential on the coordinate z. In general, the decay of the dipole lattice potential depends on the coverage Θ. Indeed, if the distance between nearest neighbors in a submonolayer of adatoms is equal to a (here $\Theta = d^2/a^2$, where, as before, d is the distance between nearest neighbors in the monolayer), the decay of the potential is proportional to $\exp(-gz)$, where $g = (2\pi/a)$ is the smallest-magnitude reciprocal lattice vector. The equality $\lambda = 2\gamma$ thus goes over to $(\pi/a) = \gamma$. Assuming that $\gamma \cong 1$ Å^{-1} (Ref. 1), we find that
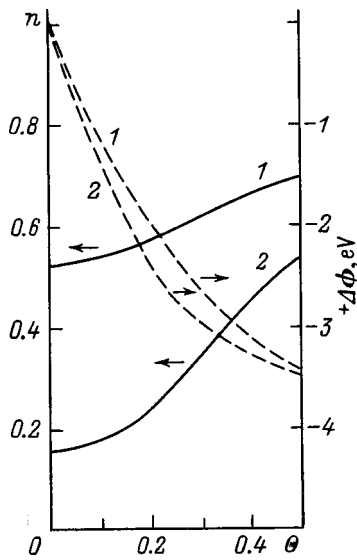
FIG. 1. Dependence of the adatom occupation numbers n (solid curves) and decrement of the work function $\Delta\phi$ (eV) (dashed curves) on the coverage Θ for Li (1) and Cs (2).

the condition $\lambda \cong 2\gamma$ holds for Θ of the order of unity (see Table I). However, as we are merely trying to discern the role of the initial state of the adatoms in the formation of the charge composition of the sputtered particle flow, cutting off the dipole potential at a distance z of the order of 1 Å does not qualitatively affect the results of the calculations.

5. It follows from Table I that all the energy parameters of the model ($\Omega, \xi, \Delta_0, \Phi$) vary smoothly in the sequence Li→Cs. We therefore confine the calculations to just the beginning and end members of this sequence. We set ε_0 equal to 0.5 eV and 1 eV, which correspond to ion velocities of 7.62×10^6 cm/s and 15.24×10^6 cm/s. The lithium atoms
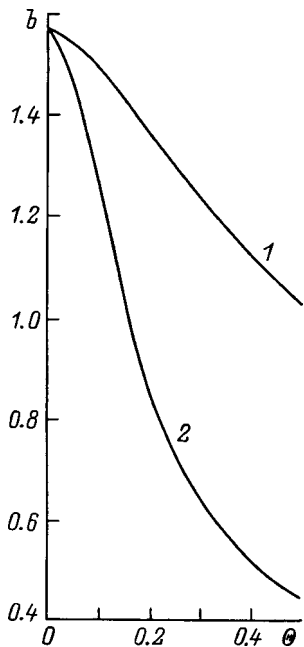
FIG. 2. Dependence of the coefficient b on the coverage. The numbering of the curves is the same as in Fig. 1.

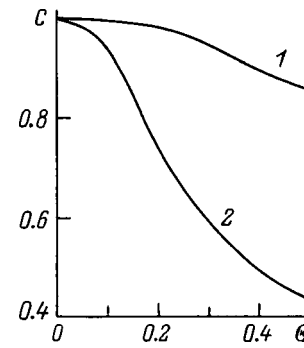
FIG. 3. The same as Fig. 2 for the coefficient c.

(ions) then have kinetic energies of 105 eV and 420 eV, and the values for cesium atoms (ions) are 2 keV and 4 keV (we assume $\gamma = 1$ Å^{-1}), respectively. These energies are typical of sputtering experiments.[15] As mentioned, the theory in Ref. 7 is valid for coverages $\Theta \leq 0.5$. These are the adatom densities investigated in our work. The results of the calculations are shown in the figures.

Figure 1 shows the occupation numbers of Li and Cs adatoms and the corresponding variations of the work function of the adsorbed system on the coverage Θ. Inasmuch as the level of a lithium adatom at zero coverage lies below the Fermi level of the substrate ($\Omega < 0$), the occupation number of a solitary lithium adatom is greater than 0.5. As Θ increases, the adatom is depolarized, and its occupation number increases. The quasi-level of a cesium adatom initially lies below the Fermi level. As the coverage increases, the quasi-level shifts downward and intersects the Fermi level of the substrate; this trend is manifested in a sharper dependence $n(\Theta)$ for Cs than for Li. The slope of the $n(\Theta)$ curve is also influenced by the narrower width of the cesium quasi-level.

Figures 2 and 3 show the coefficients b and c in Eqs. (10) and (11) as functions of the coverage. The decrease in b as Θ increases produces an increase in the occupation number $\langle n \rangle$, i.e., the emission of neutrals P^0, and a corresponding decrease in the ion escape probability P^+, as reflected in Figs. 4 and 5. It follows from a comparison of the plots of $b(\Theta)$ and $c(\Theta)$ that the ratio (c/b) increases as the coverage increases. This behavior somewhat offsets the decay of $n(\Theta)$. For $\Theta = 0$ or (as is equivalent in this case) when the
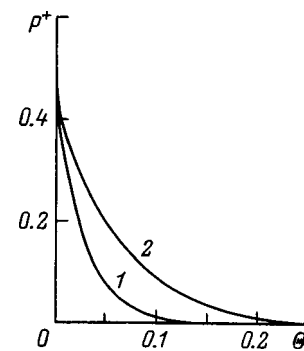

FIG. 4. Dependence of the yield P^+ of Li$^+$ ions on the coverage: (1) $\varepsilon_0 = 0.5$ eV; (2) $\varepsilon_0 = 1$ eV.
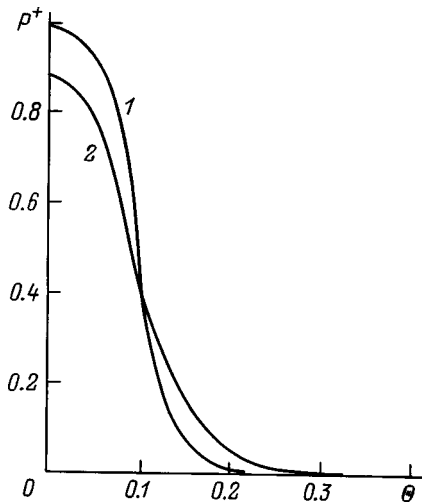
FIG. 5. The same as Fig. 4 for Cs^+ ions.

initial state of the adatoms is not taken into account, we have $b = b^* = \pi/2$ and $c = 1$. Now Eqs. (10) and (11) go over to the well-known equation (30) of Ref. 1.

The $P^+(\Theta)$ curve is steeper, the lower the ion velocity. For cesium $P^+(\Theta)$ varies from values close to unity to values close to zero, whereas for lithium it varies approximately from 0.5 to 0. This difference in $P^+(\Theta)$ for Li and Cs is associated with a corresponding difference in the positions of the quasi-levels relative to the Fermi level of the metal. The behavior of $P^+(\Theta)$ is governed mainly by the dependence of the work function on the coverage. Allowance for the initial state leads to certain quantitative corrections (see below).

Note that Eqs. (10) and (11) are asymptotic and suffer a discontinuity at the coverage determined by the equation $\bar{\Omega}(\hat{\Theta}) = 0$. The two branches of the function $n(\Theta)$ must be matched at the point $\hat{\Theta}$. Here

$$n(\hat{\Theta}) = \frac{1}{2}\left[1 + c\left(\frac{1}{b} - \frac{1}{b^*}\right)\right]. \tag{13}$$

For cesium we have $\hat{\Theta} \cong 0.11$, which gives $n(\hat{\Theta}) \cong 0.6$ and $P^+ \cong 0.4$.

Figure 6 shows the dependence of the difference $\Delta P^+ \equiv P_0^+ - P^+$ (where P_0^+ is the ion escape probability without regard for initial state of the adatoms) on the coverage for cesium. It follows from Fig. 6 that allowance for the initial-state has the effect of lowering the escape probability of
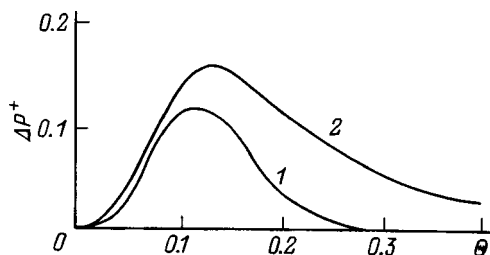
alkali metal ions (and increasing the emission of neutrals), where ΔP^+ is greater the higher the velocity of the sputtered ion. For lithium the $\Delta P^+(\Theta)$ curves are similar, but the difference $\Delta P^+ \cong 0.01$.

The range of adatom densities $\Delta\Theta$ in which the neutral and ion escape probabilities vary significantly is given by the inequality

$$-1 < \frac{\Omega + \Delta\phi}{\varepsilon_0} < 1, \tag{14}$$

i.e.,

$$\Delta\Theta \equiv \Theta_1 - \Theta_2 = \frac{\varepsilon_0}{\Phi}\left[\frac{1}{1 - n(\Theta_1)} - \frac{1}{1 - n(\Theta_2)}\right]. \tag{15}$$

Consequently, the initial state of the adatoms is what determines the width of the range of variation $P^0(\Theta)$ and $P^+(\Theta)$.

In the present study we have ignored the Coulomb shift of the adatom quasi-level

$$\delta\Omega = e^2/4l \tag{16}$$

due to the image forces. For a lithium coating $\delta\Omega \cong 3.19$ eV, which takes the quasi-level of the Li adatom from a position below the Fermi level of tungsten to a position above it. For lithium, then, the dependence of the escape probability of positive ions on the coverage $P^+(\Theta)$ is analogous to the dependence for cesium (Fig. 5). The Coulomb shift for the quasi-level of a cesium adatom is $\delta\Omega \cong 1.61$ eV, which does not qualitatively alter the dependence $P^+(\Theta)$.

In conclusion, we have estimated the role of the initial state in the formation of the charge composition of a flow of sputtered atoms. We have found that allowance for the initial state is important for the adsorption of alkali metals on the W (110) surface at a certain density of adatoms. For cesium, for example, allowance for the initial state lowers the ion escape probability, on the average, by 15% (Fig. 6), which is substantial.

[1] R. Brako and D. M. Newns, Surf. Sci. Rep. Prog. Phys. **52**, 655 (1989).
[2] A. J. Nourtier, J. Phys. (Paris) **50**, 311 (1989).
[3] J. Los and J. J. C. Geerling, Phys. Rep. **190**, 133 (1990).
[4] N. D. Lang, Phys. Rev. B **27**, 2019 (1983).
[5] R. Brako and D. M. Newns, Surf. Sci. **108**, 253 (1981).
[6] W. L. Clinton and S. Pal, Phys. Rev. B **41**, 15 (1990).
[7] J. P. Muscat and D. M. Newns, J. Phys. C **7**, 2630 (1974).
[8] I. S. Grigor'ev and É. Z. Meĭlikhov (Eds.), *Handbook of Physical Quantities* [in Russian], Énergoatomizdat, Moscow (1991), 1232 pp.
[9] C. Kittel, *Introduction to Solid State Physics*, 5th ed. [Wiley, New York (1976); Nauka, Moscow, 1978, 729 pp.].
[10] K. W. Sulston, A. T. Amos, and S. G. Davison, Prog. Surf. Sci. **26**, 63 (1987).
[11] W. A. Harrison, Phys. Rev. B **27**, 2273 (1983).
[12] K. W. Sulston, A. T. Amos, and S. G. Davison, Surf. Sci. **224**, 543 (1989).
[13] H. Nakanishi, H. Kasai, and A. Okiji, Surf. Sci. **197**, 515 (1987).
[14] H. Nakanishi, H. Kasai, and A. Okiji, Surf. Sci. **216**, 249 (1989).
[15] I. A. Baranov, Yu. V. Martynenko, S. O. Tsepelevich, and Yu. N. Yavlinskiĭ, Usp. Fiz. Nauk **156**, 477 (1988) [Sov. Phys. Usp. **31**, 1015 (1988)].

FIG. 6. Difference ΔP^+ in the yield of Cs^+ ions with (P^+) and without (P_0^+) allowance for the initial state of the adatoms as a function of the coverage: (*1*) $\varepsilon_0 = 0.5$ eV; (*2*) $\varepsilon_0 = 1$ eV.

Heating and acceleration of plasma during explosion of a wire in vacuum in a strong longitudinal magnetic field

Yu. É. Adam'yan, V. M. Vasilevskiĭ, S. N. Kolgatin, and G. A. Shneerson

St. Petersburg State Technical University, 195251 St. Petersburg, Russia
(Submitted February 2, 1998)
Zh. Tekh. Fiz. **69**, 121–127 (May 1999)

Additional estimates {Yu. É. Adam'yan, A. N. Berezkin, and G. A. Shneerson, *Abstracts of the VI All-Union Conference on the Physics of High-Temperature Plasmas* [in Russian], Leningrad (1983), Vol. 2, pp. 387–389; Yu. É. Adam'yan, V. M. Vasilevskiĭ, Krivosheev *et al.*, Pis'ma Zh. Tekh. Fiz. **21**(23), 43 (1995) [Tech. Phys. Lett. **21**(12), 968 (1995)]; S. N. Kolgatin and G. A. Shneerson, Zh. Tekh. Fiz. **67** (1), 57 (1997) [Tech. Phys. **42**, 39 (1997)]; Yu. É. Adam'yan, V. M. Vasilevskiĭ, Krivosheev *et al.*, *Abstracts of the VI International Conference on the Generation of Megagauss Magnetic Fields and Related Experiments* [in Russian], Sarov (1996)]} are presented to illustrate the possibility of and to explain the phenomenon of the acceleration and heating of a conducting medium as it expands perpendicular to a specified external field. The bulk of this paper is devoted to a description of experiments where this effect was observed during expansion of a cloud of gas with a relatively low density formed near a wire as it undergoes an electrical explosion in vacuum. The experimental data, obtained using additional diagnostics including streak camera photography and diamagnetic measurements, show that in magnetic fields with an induction of about 50 T, heating and acceleration of the outer layers of the plasma from the electrical explosion of a wire in vacuum to velocities of about 20 km/s are observed. A numerical simulation of the acceleration process is used for interpreting the results. © *1999 American Institute of Physics.* [S1063-7842(99)02205-9]

1. INITIAL ASSUMPTIONS AND MODEL PROBLEMS

The possibility of accelerating a conducting medium and heating it intensely in the presence of a rather high axial field has been demonstrated before.[3,4] The process proceeds according to the following scheme: a conducting medium (plasma or liquid conductor) is heated by a current induced in it as it moves perpendicular to the field. If the field is strong enough, the temperature rise is accompanied by an increased rate of thermal expansion, which, in turn, leads to a rise in the induced current. Under certain conditions, the process has a tendency to build up, despite the opposing effect of the electromagnetic forces. Here the acceleration and heating of the medium are at the expense of the energy of the external magnetic field. There is some interest in confirming this effect with some simple models as an example.

The first is a model of a stationary flow of conducting gas perpendicular to a field. In terms of the model problem we shall assume that the elements of the medium move along the x axis at a velocity $u_x(x)$ while all the characteristic parameters (density ρ, velocity u, pressure p, temperature T) are constant in time. Boundary values for these parameters, ρ_0, u_0, p_0, and T_0, are specified at the point $x=0$. The magnetic field B_y is equal to zero to the left of $x=0$ and rises discontinuously to B_0 at the boundary of this region and remains constant beyond it. The stationary flow of a medium with constant conductivity σ is described by the following system of magnetohydrodynamics equations:

$$\frac{d}{dx}(\rho u) = 0, \tag{1}$$

$$\rho u \frac{du}{dx} = -\frac{dp}{dx} - \delta B, \tag{2}$$

and

$$u\left(\frac{dp}{dx} - \frac{\gamma p}{\rho}\frac{d\rho}{dx}\right) = (\gamma-1)\frac{\delta^2}{\sigma}, \tag{3}$$

where $\delta \equiv \delta_z$ is the induced current density and γ is the adiabatic exponent.

It may be assumed further that the field strength in Euler coordinates is $E_y = 0$, which corresponds to the case of a plasma flow limited by conducting plates short circuited together. This implies that $\delta_z/\sigma = (-\mathbf{u}\times\mathbf{B})_z = uB_0$ or $\delta = \sigma u B_0$. Equation (1) implies that $\rho u = \rho_0 u_0$. We can then eliminate the density from Eqs.(1)–(3) and reduce them to the following system of two equations for the dimensionless velocity $U=u/u_0$ and the dimensionless pressure $P=p/p_0$, which are functions of the dimensionless coordinate $\chi = x/x_0$. The characteristic length x_0 is defined by the expression $x_0 = p_0/B^2\sigma u_0$.

The equations for the variables U and P are

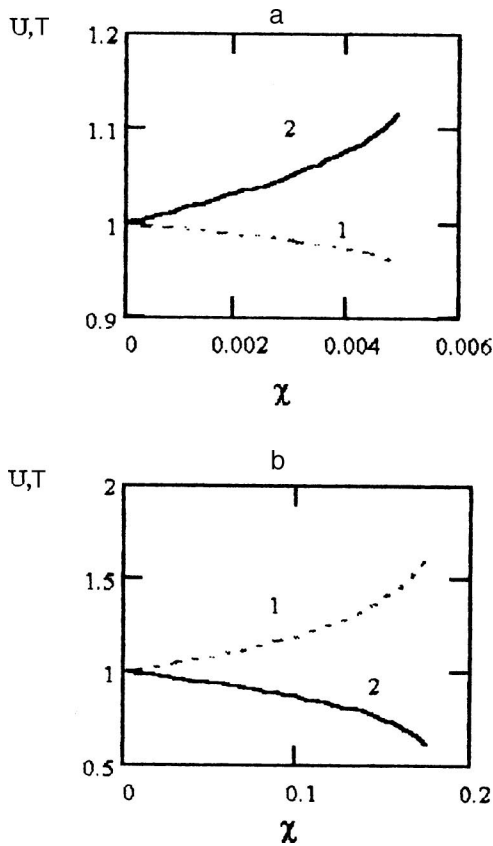$$\frac{dU}{d\chi} = \frac{U^2}{P - \dfrac{\Theta U}{\gamma}} \tag{4}$$

FIG. 1. Particle velocity and temperature in a plasma flow as it propagates perpendicular to a magnetic field (stationary flow model):(*1*) U/U_0, (*2*) T/T_0.

and

$$\frac{dP}{d\chi} = -\frac{U\left(\Theta U \dfrac{\gamma-1}{\gamma} + P\right)}{P - \dfrac{\Theta U}{\gamma}}. \qquad (5)$$

Here the characteristic parameter $\Theta = \rho_0 u_0^2/p_0$ shows up. Numerical calculations for the boundary conditions $P(0) = U(0) = 1$ show that the form of the solution is determined by the value of the parameter Θ (Fig. 1). If the initial velocity of the medium is less than the adiabatic speed of sound at the boundary (in this case $\Theta > \gamma$), then the velocity is damped with increasing x while the pressure rises. The case $\Theta < \gamma$ is more interesting; here the velocity increases with increasing x and reaches a definite value at the boundary, where the denominators on the right of Eqs. (4) and (5) change sign. Acceleration takes place over a segment whose length depends on Θ, while the velocity on the boundary of this segment equals the local adiabatic speed of sound. These remarks are confirmed by the examples shown in Fig. 1a and Fig. 1b for $\gamma = 5/3$. In the first (a), we have taken $\Theta = 2.0$ and in the second (b), $\Theta = 0.5$.

In the most interesting case of a subsonic flow, the induced currents which develop as the plasma enters the field cause it to be heated and, therefore, lead to the appearance of an additional accelerating force owing to the pressure gradi-

ent. For a subsonic flow this force exceeds the drag force, so that the velocity increases over a segment whose length is of order x_0.

The heating and acceleration of the plasma develop over a length of order x_0 over a time

$$\tau_0 \approx \frac{x_0}{u_0} \approx \frac{\rho_0}{B^2 \sigma \Theta}. \qquad (6)$$

To within a factor of $1/\Theta$, τ_0 is a parameter that is known in magnetohydrodynamics, the inductive drag time.[5] For $\Theta > \gamma$ the plasma is actually slowed down over a time of order τ_0. However, for $\Theta < \gamma$ the significance of this parameter changes to the opposite and τ_0 becomes the time scale over which the plasma accelerates as it moves across the field. For a copper vapor density on the order of 1 kg/m^3, which corresponds to a density of 10^{25} m^{-3}, accelerating a plasma with a conductivity of 10^3 S/m over a time on the order of 10^{-6} s requires that it expand across a field with an inductance on the order of 10^2 T. The threshold value of the inductance is greater for higher densities and lower conductivities.[3]

These qualitative features of plasma heating in a strong field are confirmed by a series of calculations using the full model of a nonstationary radial magnetohydrodynamic flow for a hypothetical medium with a fixed conductivity. The same computer code was used as in Ref. 3. In the model problem it was assumed that the plasma cylinder begins to expand in an axial field B_0 from an initial state characterized by a radius r_0, temperature T_0, and density ρ_0. As noted, the conductivity was assumed constant. In the calculations interpolated equations of state were used.[6] As the series of curves in Fig. 2a shows, the temperature of the outer layers of a plasma without a field decreases in time owing to adiabatic expansion (curve *1*), but for a conductivity of 10^4 S/m and $\rho = 300$ kg/m^3, by the time a field with an inductance of 5 T is applied, the temperature rises rapidly following a drop (curve *2*). Further increases in the external field lead to more rapid heating (curve *3* for 10 T, curve *4* for 20 T, and curve *5* for 50 T). There is no heating in a 50 T field at this density if the conductivity falls to 4×10^2 S/m (the corresponding temperature variation for the outer plasma layers is shown in curve *1*) but heating occurs if, along with a drop in the conductivity to this value, the initial density is reduced to 3 kg/m^3. (The heating process for this case is illustrated by Fig. 2b.) In electrical explosions of wires, the process is more complicated than in these model problems in that the conductivity depends on the temperature and density.

Previous calculations[2,3] have shown that plasma heating by an induced current and acceleration of the boundary only take place in fields above 10^2 T, if the initial heating of the conductor is determined by a current with a density on the order of $10^{10} - 10^{11}$ A/m^2. The threshold field can be substantially lower for the low density plasmas which are usually formed near a wire as it explodes in a vacuum.

If we assume that the particle temperature resulting from breakdown of a gas and its ionization is on the order of 1 eV, then this corresponds to a thermal velocity for copper atoms of $\sim 10^3$ m/s. Over a time on the order of 1 μs, the cloud of
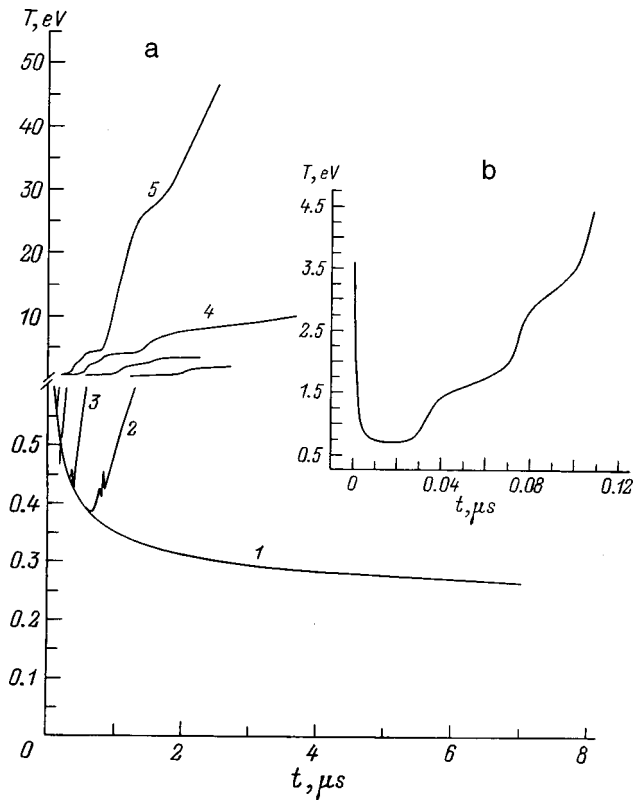
FIG. 2. Time dependence of the temperature at the edge of the plasma cylinder as it expands in a longitudinal magnetic field. Initial radius $r_0 = 1$ mm, $T_0 = 3$ eV.

FIG. 3. Single turn solenoid with an evacuable cylinder containing a wire inside it.

gas moves a distance of roughly 1 mm, while the concentration of particles in this zone is on the order of 10^{24} m^{-3}.[7] According to these estimates, for a conductivity $\sigma = 100$ S/m, we might expect acceleration of the plasma boundary in a field $B_z = 50$ T. The experiments described below confirm this hypothesis.

DESCRIPTION OF THE EXPERIMENTAL APPARATUS

In these experiments a copper wire with a diameter of 0.2 mm and length of 15 mm was placed in an evacuated glass cylinder *1* (Fig. 3) with an inner diameter of 16 mm. During fabrication, the wire was outgassed for a day at a temperature of 400 °C and a pressure of 0.1 Pa. The cylinder was placed in the cavity of a single-turn solenoid with a replaceable destructible insert. The solenoid was powered by a low inductance capacitor bank with a capacitance of 648 μF charged to 35 kV. The maximum inductance was 70 T. The wire was exploded near the peak inductance by discharging an additional capacitor bank. The rise time of the current in the wire to its peak (1 μs) was considerably shorter than the field rise time (7 μs). MHD calculations confirm that under these conditions the longitudinal field can be regarded as quasistationary from the standpoint of its effect on the electrical explosion of the wire. The design of the current lead to the cylinder with the exploding wire was in the form of a ''squirrel cage'' of thin wires in order to eliminate any mechanical effect of the longitudinal magnetic field
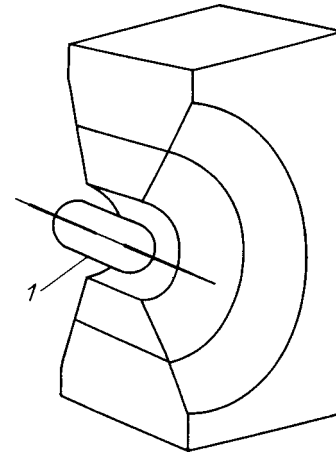
on the elements of the discharge circuit. The circuit parametersfor the exploding wire were capacitance 3.75 μF, inductance 300 nH, and charging voltage 10 kV.

The current in the wire was measured with a Rogowsky loop and the voltage on the exploding wire, with a voltage divider with a compensating inductive component. In order to clarify how the longitudinal magnetic field affected the explosion process, these methods were supplemented by two additional channels: streak camera photography of the plasma light, which made it possible to obtain the radial distribution of the plasma light of the electrical explosion at all times and diamagnetic signal measurements.

Streak camera photography of the plasma light makes it possible to measure the velocity at which the plasma boundary expands. Figure 4 shows a sketch of the mutual locations of the solenoid, exploding wire, and streak camera. The triggering of the streak camera scan was synchronized with the time of the trigger pulse to the spark gap in the exploding wire circuit.

The electrical conductivity of the plasma for a known expansion velocity was determined by measuring the diamagnetic signal owing to the magnetic flux squeezed out of the plasma. The essence of the diamagnetic measurements is clarified in Figs. 5a and 5b. Two windings were located on the surface of the vacuum cylinder containing the exploding wire; one of them (the measurement loop) lay in a plane passing through the center of the wire and the second (the compensating loop) lay in the region of the current lead. At the start of a shot the magnetic field of the external source rises over a characteristic time much longer than the time for the field to diffuse into the exploding wire, so that at the time ofthe explosion, which occurs near the maximum in B_z, the field inside the diamagnetic loop is essentially uniform along the radius. Then, because the compensating loop is inserted with a polarity opposite that of the measurement loop, the signal at the output of the integrator is small. The expanding plasma from the electrical explosion squeezes the magnetic flux out of the cross section of the measurement loop because of the azimuthal plasma current flow generated by the radial motion of the conducting medium. Estimates show
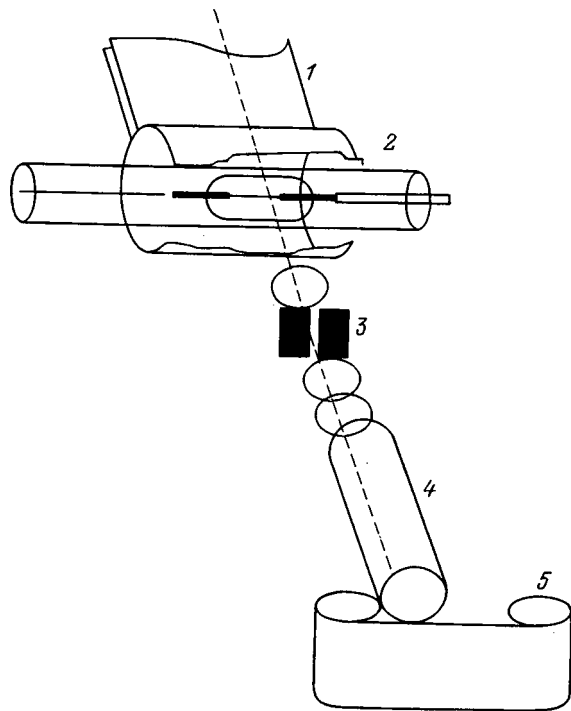
FIG. 4. A sketch showing the locations of the exploding wire, solenoid, and streak image tube: (*1*) nominal view of the external field solenoid, (*2*) exploding wire with current leads, (*3*) slit diaphragm, (*4*) electron-optical image tube, (*5*) photographic film.

that the component of the azimuthal current owing to the nonstationary character of the external field does not have a significant effect on the explosion process under these experimental conditions.

2. EXPERIMENTAL RESULTS

Figure 6 shows typical oscilloscope traces of the current and voltage in an exploding wire experiment for $B_z = 0$ and 50 T. For the given circuit parameters the explosion has practically no effect on the trace of the current through the wire. The effect of the field shows upas a sharp drop in the amplitude of the voltage across the exploding wire, which indicates a rise in the plasma conductivity. Pictures of exploding wires with and without a longitudinal magnetic field are shown in Fig. 7.

Figure 8 shows the results of analyzing the streak camera pictures of an explosion; the curves represent the contours of equal blackening of the photographic plate. The picture in Fig. 7a corresponds to an explosion in a longitudinal magnetic field and is divided int two distinct zones with differing brightnesses. The outer zone is a rapidly expanding, low density plasma. The expansion of the plasma is limited by the walls of the vacuum vessel. The inner zone is much brighter and expands slowly, almost the same as in an electrical explosion without a longitudinal magnetic field (Fig. 7b). Thus, it is clear from a comparison of the streak photographs that for the parameters of the exploding wire circuit in these experiments, electrical explosions in a longitudinal magnetic field differ in having a rapidly expanding, low density plasma shell. This is confirmed by comparing the contours of equal blackening in the pictures with and without a
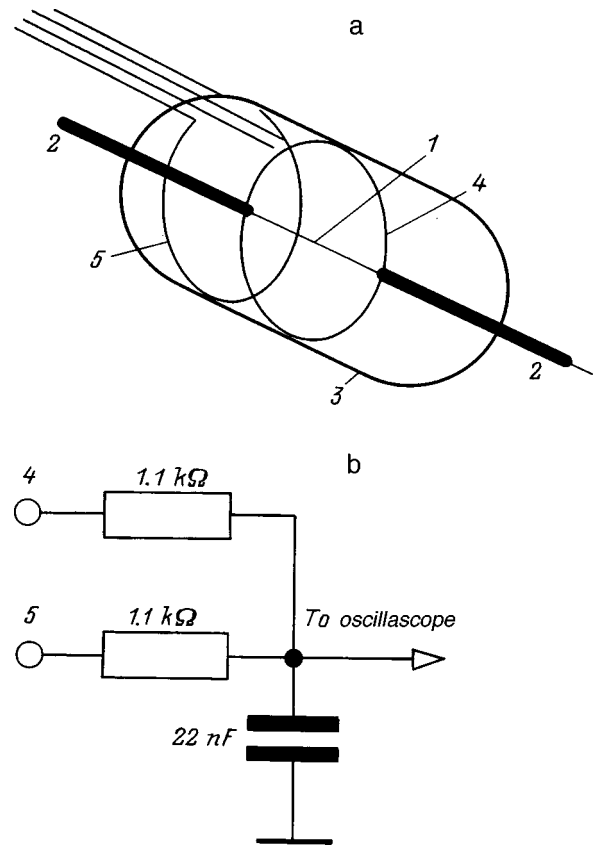

FIG. 5. a: A sketch of the location of the diamagnetic loops on the surface of an evacuated cylinder containing the exploding wire: (*1*) exploding wire, (*2*) current leads, (*3*) glass cylinder, (*4*) measurement loop, (*5*) compensating loop. b: Sketch of the summation–integration circuit for measuring the magnetic flux expelled by the plasma of an electrically exploding wire.

field (Fig. 8). The profiles corresponding to densities of 100 and 7% (for the experiment with a field) are also indicated right on the streak camera pictures.

The distribution of the relative local radiant intensity of the plasma was obtained by processing the photometric scans of the stream camera pictures assuming a cylindrical emitting region that was transparent with a low density. An analytic solution of the Abel equation, i.e., Abel inversion,[7] was used in processing the data:
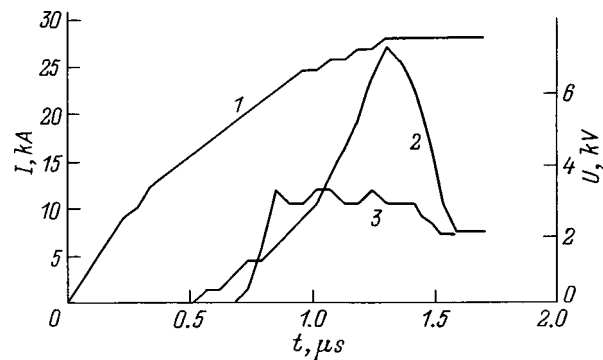

FIG. 6. Oscilloscope traces of the currents and voltages during the electrical explosion of a wire with and without a longitudinal magnetic field: (*1*) current, (*2*) voltage for $B = 0$, (*3*) voltage for $B = 50$ T.
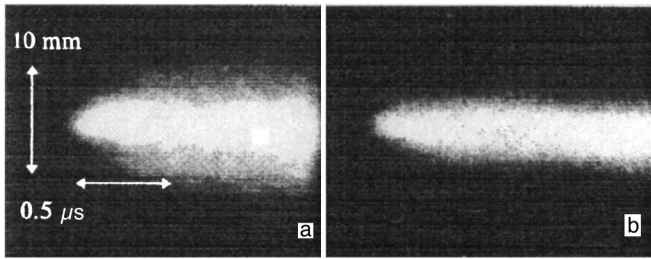
FIG. 7. Streak camera pictures of electrical explosions of wires in a longitudinal magnetic field $B_z = 50$ T (a) and without a field (b).

$$I(r) = \frac{1}{\pi} \int_r^a \frac{dJ(y)}{dy} \frac{1}{\sqrt{y^2 - r^2}} dy. \tag{7}$$

Here $J(y)$ is the distribution of the film blackening along the spatial coordinate y; r is the instantaneous radius of the point at which the source intensity $I(r)$ is being determined; and, a is the outer radius of the plasma column at the given time. In solving Eq. (1) the values of $J(y)$ obtained from photometric scans of the streak camera pictures were approximated by polynomials. The innermost layers, which are obviously not optically transparent were excluded from the analysis.

The results of this analysis are shown in Fig. 9. The show that the expansion velocity of the plasma boundary is close to 20 km/s, which corresponds to a copper ion energy on the order of 100 eV.

An oscilloscope trace of the diamagnetic signal for $B_z = 50$ T is shown in Fig. 10. There one can distinguish a signal proportional to the external magnetic field, which is present in this trace because it has not been completely compensated, as well as a signal proportional to the expelled magnetic flux, which stands out easily against the background of the residual uncompensated signal.
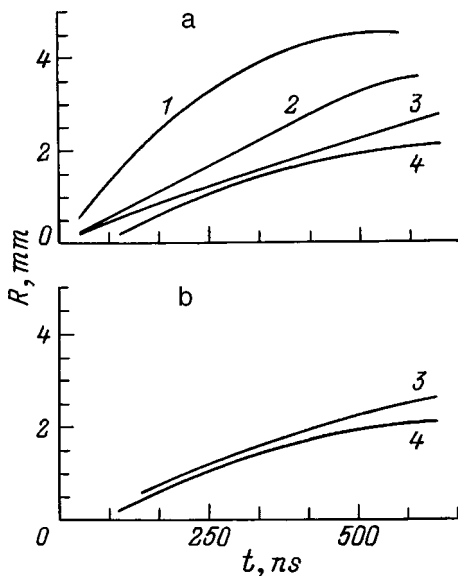

FIG. 8. Curves corresponding to equal blackening of the streak camera photographic plate: (*1*) 7, (*2*) 20, (*3*) 35, (*4*)100; $B_z = 50$ T (a), 0 (b).
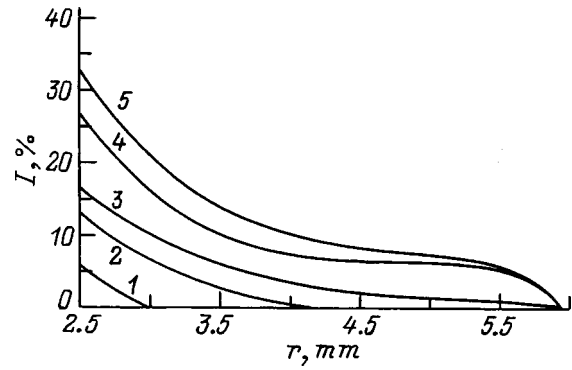

FIG. 9. Radial distribution of the relative local luminous intensity of the outer region of the plasma obtained by processing streak camera pictures for $B_z = 50$ T. I is normalized to the intensity of the central region. Time t (ns): 50 (*1*), 150 (*2*), 250 (*3*), 350 (*4*), and 450 (*5*).

DATA ANALYSIS

Figure 11 shows the time dependence of the resistance of the wire in some experiments with a field. Over times $t_1 < t < t_2$, when the voltage on the wire changes little, the resistance varies between 0.14 and 0.1 Ω. For comparison, we note that in an experiment without a field the resistance at time t_2 is close to 0.2 Ω. Judging from the streak camera pictures, the brightly luminous region differs little in shots with and without a field. We can assume, therefore, that the resistance of the high density region is roughly the same in both shots. The measured resistances of the channel differ because in the shots with a magnetic field the dense plasma region is shunted by the low density region. At time t_2, when the phase transition is completed, the resistance of the outer zone of the discharge is about 0.17 Ω under these assumptions. Then we can estimate the average conductivity, since by t_2 the plasma fills up the entire cylinder. The conductivities about 400 S/m.

The diamagnetic signal together with the measurements of the expansion velocity of the plasma can be used to estimate the plasma conductivity in another way. The simplest estimate of the conductivity is based on the following considerations: $\Delta\Phi \approx \Delta B \cdot \pi R^2$, where $\Delta\Phi$ is the magnetic flux expelled from the plasma through the time it comes into contact with the walls and R is the vessel radius. Since ΔB
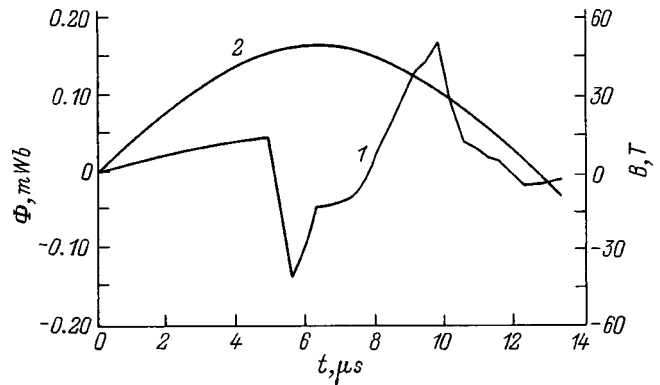

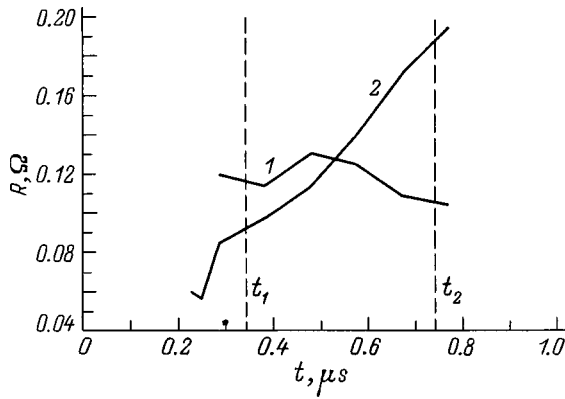FIG. 10. Oscilloscope traces of the diamagnetic signal (*1*) and longitudinal magnetic field (*2*).

FIG. 11. Time variation of the resistance of the exploding wire for $B = 50$ T (*1*) and 0 (*2*).

$\approx i'_\varphi \mu_0$, where $i'_\varphi \approx u \sigma B R$ is the linear density of the induced current and u is the expansion speed of the plasma boundary, for $B = 50$ T, $u = 20$ km/s, and $R = 8$ mm, we obtain σ approx. 100 S/m. For a more exact estimate of σ we calculated the diffusion of the magnetic field into a cylindrically expanding plasma with a specified conductivity. The diffusion equation for the magnetic field in cylindrical coordinates,

$$\frac{\partial B_z}{\partial t} = \frac{\partial}{r \partial r}\left(\chi r \frac{\partial B_z}{\partial r}\right) - B_z \frac{\partial(ru)}{r \partial r}, \qquad (8)$$

was solved numerically, where $\chi = 1/(\mu_0 \sigma)$ is the diffusion coefficient for the electromagnetic field, σ is the electrical conductivity of the plasma, μ_0 is the magnetic permeability of the medium, B_z is the induction of the longitudinal magnetic field, and u is the plasma expansion velocity. It was assumed that the conductivity is constant over the plasma cross section, while u increases linearly from the axis of the wire to the boundary, which corresponds to uniform expansion of the column. Figure 12 shows the diamagnetic signal calculated in this way at the time of the maximum plasma radius as a function of the conductivity.

The measured diamagnetic signal corresponds to the calculated value for $\sigma \approx 200$ S/m. The drop and polarity reversal of the diamagnetic signal after it reaches a maximum can be explained by the stopping of the outer plasma layers when they collide with the walls of the cylinder, after which there is a "dip" in the longitudinal magnetic field inside the
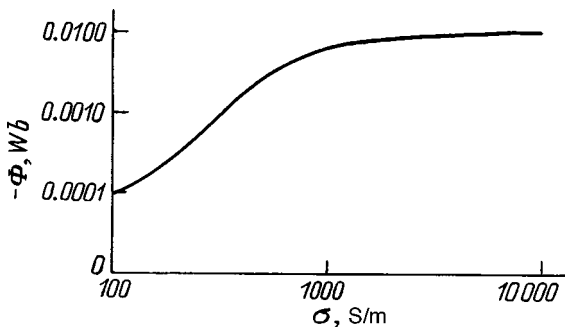

FIG. 12. Calculated diamagnetic signal as a function of the plasma conductivity for an expansion velocity $u = 20$ km/s and radius $r = 8$ mm of the plasma boundary.

plasma. In addition, the plasma may be reflected from the walls, and this may cause pinching of the plasma and trapping of flux. Here we have only considered the initial part of the diamagnetic signal corresponding to expansion of the plasma.

In the next stage of the calculations we solved a complete system of MHD equations for the radial flow analogous to that described in Ref. 3 with the temperature and density dependence of the conductivity taken from Ref. 8. Here the initial values of the density and temperature were varied. The calculations yielded a diamagnetic signal that differed substantially from the measured signal. Thus, we made a series of calculations with a lower conductivity, which, as in the calculations with a linear radial dependence of the velocity, was taken to be constant at $200-400$ S/m. A similar series of calculations was done for different values of the initial density and temperature. The calculations showed that for an initial density below ~ 5 kg/m^3 and temperature above ~ 3 eV, the plasma enters a regime of acceleration, accompanied by a rise in the temperature of the outer layers to about 100 eV. The resulting velocity distribution differs greatly from the linear distribution proposed for preliminary estimates of the plasma conductivity based on the diamagnetic measurements. The best agreement with the experimental data for $\sigma \approx 400$ S/m was obtained for the calculation with an initial density of ~ 1 kg/m^3 and an initial temperature ~ 5 eV.

Since radiative losses were neglected in the calculations, the temperature results are estimates. The most reliable result of the test calculations is that for all variants of the initial conditions, a diamagnetic signal close to the measured value could be obtained only for conductivities on the order of $\sim 10^2$ S/m, which corresponds to the estimates based on the resistance of the channel. This estimate is more than an order of magnitude smaller than conventional tabulated values.[8] The discrepancy may be an indication that the formulas for the conductivity based on the assumption of thermodynamic equilibrium of the medium are no longer valid at low densities, where calculations show that the particle density is approximately 10^{22} m^{-3}. Since the azimuthal current density is $\sim 10^8$ A/m^2, for these electron densities their drift velocity is on the order of 10^5 m/s, which is substantially higher than the ion thermal speed. Under these conditions, the assumption of a lack of thermodynamic equilibrium and a drop in the conductivity owing to the excitation of chaotic oscillations in the plasma appears to be fully justified.

With the above estimates of the conductivity and expansion velocity of the boundary of the wire it is easy to calculate the energy applied to the plasma through heating by the azimuthal current over a time on the order of 10^{-6} s. If the conductivity is a few times 10^2 S/m, then this energy is 10^7 J/m^3, while the current density is $\sim 10^8$ A/m^2. This estimate confirms the feasibility of using electrical explosions in very high fields for plasma heating.

CONCLUSION

1. It has been shown experimentally that a strong axial magnetic field has a substantial effect on the electrical explosion of thin wires in vacuum. Changes in oscilloscope traces

of the current and voltage across an exploding wire have been observed when an external longitudinal field is applied.

2. Model problems demonstrate the possibility of self excitation of an azimuthal current and plasma acceleration during thermal expansion in a strong magnetic field. A possible buildup of this process has been noticed, where the heating increases with rising azimuthal current, which, in turn, leads to acceleration of the plasma and a further current rise. It has been shown that the process has a threshold character.

3. The experiments confirmed the possibility of plasma heating by the induced azimuthal current and the acceleration of the plasma. A sudden expansion of the relatively low-density, cylindrical plasma cloud formed during the explosion of a wire in a vacuum was observed. It has been shown that in a 50 T field the velocity of the boundary approaches $\sim 2 \times 10^4$ m/s.

4. Measurements of the expelled magnetic flux confirmed the existence of an induced azimuthal current with a density on the order of 10^8 A/m^2 and an expanding plasma with a conductivity close to 10^2 S/m.

We thank S. V. Bobashov for useful discussions of these results.

[1] Yu. É. Adam'yan, A. N. Berezkin, and G. A. Shneerson, *Abstracts of the VI All-Union Conference on the Physics of High-Temperature Plasmas* [in Russian], Leningrad (1983), Vol. 2, pp. 387–389.

[2] Yu. É. Adam'yan, V. M. Vasilevskiĭ, Krivosheev *et al.*, Pis'ma Zh. Tekh. Fiz. **21**(23), 43 (1995) [Tech. Phys. Lett. **21**(12), 968 (1995)].

[3] S. N. Kolgatin and G. A. Shneerson, Zh. Tekh. Fiz. **67**(1), 57 (1997) [Tech. Phys. **42**, 39 (1997)].

[4] Yu. É. Adam'yan, V. M. Vasilevskiĭ, Krivosheev *et al.*, *Abstracts of the VI International Conference on the Generation of Megagauss Magnetic Fields and Related Experiments* [in Russian], Sarov (1996).

[5] T. G. Cowling, *Magnetohydrodynamics* [Hilger, London(1975); Atomizdat, Moscow (1978)].

[6] S. N. Kolgatin and A. V. Khachatur'yants, Teplofiz. Vys. Temp. **78**, 447 (1982).

[7] G. S. Sarkisov, B. Étlisher, V. V. Yan'kov *et al.*, Zh. Éksp. Teor. Fiz. **108**, 1355 (1995) [JETP **81**, 743 (1995)].

[8] I. M. Bespalov and A. Ya. Polishchuk, IVT AN SSSR Preprint No. 1-257 [in Russian], Institute of High Temperatures, USSR Academy of Sciences, Moscow (1988), 35 pp.

Translated by D. H. McNeill

BRIEF COMMUNICATIONS

Fractal model for polarization switching kinetics in ferroelectric crystals

R. P. Meĭlanov and S. A. Sadykov

Dagestan State University, 367025 Makhachkala, Russia
(Submitted October 19, 1998)
Zh. Tekh. Fiz. **69**, 128–129 (May 1999)

The kinetics of polarization switching in ferroelectric crystals is studied using the concept of fractals. The switching process is found to be multifractal. © *1999 American Institute of Physics.*
[S1063-7842(99)02305-3]

In this paper we propose a method for analyzing the kinetics of switching and the shape of the transition current in terms of the concept of fractals, which is based on the concept of a space–time metric of fractional dimensionality.

The most widespread models for analyzing the kinetics of polarization reversal currents are based on the Kolmogorov–Avrami (K–A) statistical theory of crystallization.[1,2] Of these the most rigorous are the model theories of Fatuzzo[3] and Ishibashi and Takagi.[4] In the modified K-A theories[5–9] calculating the switching current reduces to finding the fraction of the switched volume, $Q(t)$, out of the total volume of a ferroelectric capacitor. Then the switching current is given by

$$i(t) = 2P_s A \frac{dQ(t)}{dt}, \qquad (1)$$

where P_s and A are the spontaneous polarization and the area of the sample electrode.

These models are based on simplifications of the domain structure dynamics and they also neglect such important factors as internal screening, stochastic changes in the dielectric properties, the effects of anisotropy and delay of the interactions, and the features of the electronic spectrum of the system. As a result, agreement with experiment is obtained only for fractional dimensionalities of the kinetic process and it has not yet been possible to approximate the falling part of the current pulse satisfactorily.

In terms of a number of their dynamic characteristics, ferroelectric crystals behave as fractal systems. It is known that polarization reversal takes place through the appearance of a certain number of seeds for domains of the opposite polarization which subsequently grow in the forward and lateral directions. According to the Fatuzzo-Mertz-Miller-Weinrach- Hayashi model,[10] the end and side displacements of the domain walls proceed through wall seed formation, i.e., the development of seeds on the domain walls. Later, seeds can be formed on growing steps which develop before the seeds. The appearance of structures of this type can be regarded as a sort of intermediate state of the material during a transition from one polar state to the other. The system manifests fractal properties, since lateral growth of the do-main walls is the result of the formation of self-similar structures (seed–steps).

Note that fractal analysis goes beyond the geometric version. The spatial disorder of a system leads to a power law dependence for the temporal relaxation of nonequilibrium states when memory effects are present. Thus, in a study[11] of the dynamics of prephase states in polydisperse ferroelectrics, it has been shown that the polydispersity index is a function of the fractal dimensionality and that this index determines the time evolution of the system. In Ref. 12 the equations of motion for a domain wall were modified with the time derivatives of integral order being replaced by derivatives of fractional order, so that it was possible to describe features of dielectric permittivity spectrum which are usually obtained by empirical selection of the relaxation time distribution functions.

The physical justification for using the mathematical apparatus of fractional differential equations for analyzing the properties of systems with fractal structures has been discussed elsewhere.[13–15] In particular, it has been shown[13] that using a fractional time derivative corresponds to taking into account memory effects in a system. Thus, given the importance of memory effects in polarization-reversal processes in ferroelectrics, we generalize Eq. (1) for the switching current using a fractional time derivative,

$$i(\xi) = \frac{2P_s A}{t_0} \frac{d^\alpha}{d\xi^\alpha} Q(\xi), \qquad (2)$$

where $\xi = t/t_0$, t_0 is the characteristic switching time, α is the dynamic fractal dimensionality ($0 < \alpha \leqslant 1$), and $2P_s A Q(\xi)$ is the switching charge.

For $\alpha = 1$, which corresponds to the absence of memory effects, Eq. (2) is the same as Eq. (1).

In the case where the simplest expression of the form $Q(\xi) = 1 - \exp(-\xi)$ is used for the time dependence of the polarization charge, Eq. (2) gives

$$j(\xi) = \frac{1}{\Gamma(2-\alpha)} \xi^{1-\alpha} \exp(-\xi) \, _1M_1(1-\alpha; 2-\alpha; \xi), \quad (3)$$

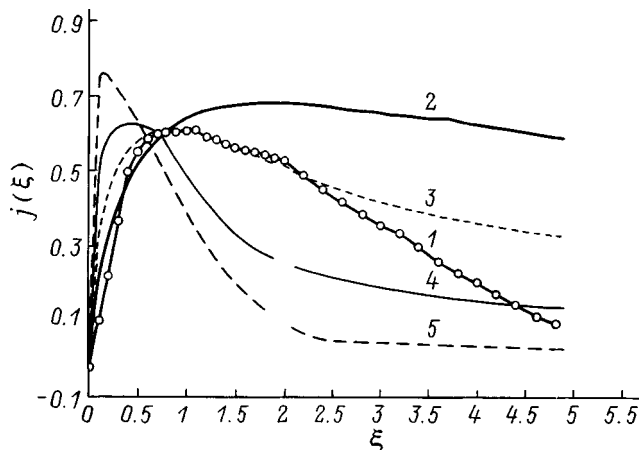where $j(\xi) = i(\xi) t_0 / (2P_s A)$ is the dimensionless transition

FIG. 1. The switching current $j(\xi)$ in a thin film of KNO_3 as a function of $\xi = t/t_0$: *1* — experiment at 55 °C, $t_0 = 185$ ns; *2–5* — calculated for $\alpha = 0.3, 0.5, 0.7,$ and 0.9, respectively.

current, $\Gamma(\alpha)$ is the Euler gamma function, and $_1M_1(a;b;z)$ is Kummer's confluent hypergeometric function.[16]

For $\alpha = 1$, Eq. (3) yields $j(\xi) = \exp(-\xi)$. This function does not correspond to the experimental data, and this was the reason for the various modifications of the expression for $Q(\xi)$ in the K-A theory.[4–8] In particular, an expression of the form $Q(\xi) = 1 - \exp(-\xi^n)$ is used, where the parameter characterizes the dimensionality of domain growth. In this case, satisfactory agreement with the experimental data is obtained for nonintegral values of n.

The results of a numerical simulation of the transition current obtained using Eq. (2) for different dynamic fractal dimensions are shown in Fig. 1, which also shows some experimental data from Ref. 5. It can be seen that even with the simplest approximation, $Q(\xi) = 1 - \exp(-\xi)$, the proposed model can describe the polarization-reversal kinetics quantitatively, with the best agreement for a multifractal process. The values vary from 0.3 in the initial stage to 0.9 in

the final stage of polarization reversal. The predominant value of the fractal dimensionality, which was determined from the ratio $\Delta t(\alpha)/T$ (where T is the total duration of the experiment and $\Delta t(\alpha)$ is the time over which the experimental data for the switching current coincide with the calculated values for a given α), is $\alpha = 0.6$.

The nature of the dynamic fractal dimension α arises from the processes involved in formation of the switching current. Clarifying the specific mechanisms for formation of α is a separate problem.

In conclusion, we note that the variation in the switching charge should also be examined using a fractal approach, which would make it possible to obtain more detailed information on the mechanism for the restructuring of the domain structure in ferroelectrics.

[1] A. I. Kolmogorov, Izv. Akad. Nauk SSSR Ser. Mat. **3**, 355 (1937).

[2] M. Avrami, J. Chem. Phys. **7**, 1103 (1939).

[3] E. Fatuzzo, Phys. Rev. **7**, 1999 (1962).

[4] Y. Ishibashi and Y. J. Takagi, J. Phys. Soc. Jpn. **31**, 506 (1971).

[5] K. Dimler, M. Parris, D. Butler *et al.*, J. Appl. Phys. **61**, 5467 (1987).

[6] J. F. Scott, L. Kammerdiner, M. Parris *et al.*, J. Appl. Phys. **64**, 787 (1988).

[7] H. Orihara and Y. Ishibashi, J. Phys. Soc. Jpn. **61**, 1919 (1992).

[8] V. Ya. Shur, E. L. Rumyantsev, and S. D. Makarov, Fiz. Tverd. Tela (St. Petersburg) **37**, 1687 (1995) [Phys. Solid State **37**, 917 (1995)].

[9] V. Ya. Shur, S. D. Makarov, N. Yu. Ponomarev *et al.*, Fiz. Tverd. Tela (St. Petersburg) **38**, 1889 (1996) [Phys. Solid State **38**, 1044 (1996)].

[10] M. Lines and A. Glass, *Principles and Applications of Ferroelectrics and Related Materials* [Clarendon, Oxford (1977); Mir, Moscow (1981), 736 pp.].

[11] Yu. G. Kubarev, Izv. Akad. Nauk SSSR, Ser. Fiz. **57**(3), 129 (1993).

[12] H. A. Galiyarova, Ferroelectrics **170**, 111 (1995).

[13] R. R. Nigmatullin, Teor. Mat. Fiz. **90**, 354 (1992).

[14] K. V. Chukbar, Zh. Éksp. Teor. Fiz. **108**, 1875 (1995) [JETP **81**, 1025 (1995)].

[15] R. P. Meĭlanov, Pis'ma Zh. Tekh. Fiz. **22**(23), 40 (1996) [Tech. Phys. Lett. **22**(12), 967 (1996)].

[16] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* [Dover, New York (1965); Nauka, Moscow (1979)].

Translated by D. H. McNeill

Effect of carbon ion bombardment and surface oxidation on the twinning rate of single-crystal bismuth

O. M. Ostrikov

Mozyr State Pedagogical Institute, 247760 Mozyr, Belarus
(Submitted March 30, 1998)
Zh. Tekh. Fiz. **69**, 130–131 (May 1999)

A study is made of the change in the dependences of the normal velocity of twinning boundaries on the magnitude of shear stresses in the twinning plane $v_n = v_n(\tau)$ in bismuth crystals owing to ion-cluster doping and oxidation of the irradiated surface. Irradiation was by 25 keV carbonions at a dose of 10^{17} ion/cm^2. Twinning of the crystals took place under pulsed loading conditions with pulse durations of 10^{-5}–10^{-4} s and stress amplitudes of $(0.2$–$2.0) \times 10^3$ g/mm^3. Carbon ion bombardment of single-crystal bismuth causes a shift in the $v_n = v_n(\tau)$ curve toward lower stresses. An oxide film slows down the motion of twinning dislocations. © *1999 American Institute of Physics.* [S1063-7842(99)02405-8]

The effect of doping, tempering, and oxidation of crystalline bismuth surfaces on their twinning rate has been studied[1] and it was found that these factors produce the same changes in the $v_n = v_n(\tau)$ curves, specifically, they shift them toward higher stresses. There is also some interest in studying the effect of ion-cluster doping on the mobility of twinning boundaries, since this form of energy interaction with crystals can be used for extensive modification of their mechanical properties,[2] including their plastic properties. The physical problem is then to study the effect of structural changes induced by anion beam in the surface layers on the twinning process in crystals.

The purpose of this paper was to study the effect of the bombardment of single-crystal bismuth by carbon ions on the twinning velocity of the crystals during impulse loading.

EXPERIMENTAL TECHNIQUE

Single crystals of bismuth with sizes of $5 \times 10 \times 70$ mm were grown by the Bridgman method from a 99.999% pure stock and had an initial density of 10^5 cm^{-2} basal and 10^3 cm^{-2} pyramidal dislocations. Prismatic samples with dimensions of $2 \times 5 \times 10$ mm were used in the experiments. One of the side facets of the sample, on which the displacement of the twinning boundaries was measured after each load pulse, coincided with the shear plane of the atoms during twinning and the end surfaces of the samples were faceted by the (111) cleavage plane. The lateral facets of the samples were chemically polished.[3] The methods used for deformation and for calculating the stresses and pulse duration have been discussed elsewhere.[4]

Bismuth was chosen as the test material because plastic deformation of bismuth by twinning is relatively easy compared to other materials (Schmid factor 0.48). In addition, the melting temperature of bismuth is a relatively low 273 °C.[5]

The samples were bombarded with carbon ions because carbon dissolves very poorly in bismuth. The tiny solubility of carbon in bismuth at its melting point is 0.0003 wt% (0.0052 at. %).[5] This property of the binary Bi–C system allows us to exclude any effect on the twinning process owing to chemical compounds formed by interactions of implanted impurity atoms with target atoms and makes modeling the mechanisms for twinning of irradiated crystals significantly easier.

In these experiments single crystals of bismuth were bombarded by 25 keV carbon ions to a dose of 10^{17} ion/cm^2. The load pulse duration was 10^{-5}–10^{-4} s and the stress amplitude was $(0.2$–$2.0) \times 10^3$ g/mm^2.

The expansion velocity of the twins (normal velocity v_n) was calculated by measuring the displacement of the twin boundaries in the shear plane after each load pulse and assuming that the boundaries move uniformly during the time the pulse is applied. Here, as a rule, several parallel interlayers of a single crystallographic system developed simultaneously in the sample. Twinning took place until neighboring interlayers collided with one another. The thickness of the interlayers reached 400–500 μm by the end of the experiment.

EXPERIMENTAL RESULTS AND DISCUSSION

The results of the measurements are shown in Figure 1. Each point in the graph is the arithmetic average of measurements on several interlayers. A comparison of data sets *1–3* shows that carbon ion bombardment of single crystals of bismuth stimulates the mobility of twinning dislocations and shifts the $v_n = v_n(\tau)$ curve toward lower stresses. Oxidation of the irradiated surface, on the other hand, slows down the development of the twinning interlayers somewhat, and shifts the $v_n = v_n(\tau)$ curves toward higher stresses than those at which twinning occurs in irradiated crystals without an oxide film.

On a semilogarithmic plot, the experimental points of the $v_n(\tau)$ curves in the figure fit a straight line well, so they can be approximated by an expression of the type
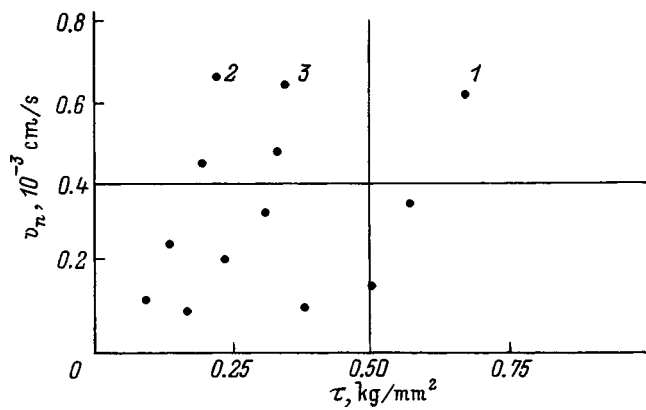
FIG. 1. Normal rate of growth of a twinning interlayer as a function of stresses: *1* — unirradiated single-crystal bismuth, *2* — single-crystal bismuth irradiated with carbon ions, *3* — irradiated single-crystal bismuth with an oxide film.

$$v_n = A \exp \frac{\gamma \tau}{kT}, \qquad (1)$$

where A is a constant, γ is the activation volume, and k and T are the Boltzmann constant and absolute temperature.

The exponential form of the $v_n(\tau)$ curves indicates that a thermally activated mechanism predominates in overcoming the barriers to twinning dislocations during twin development. The activation volume found from the slope of the $\ln v_n(\tau)$ curves was $\gamma = 34a^3$ for the unirradiated crystal (where a is the lattice parameter), and $\gamma = 56a^3$ for $\gamma = 48a^3$ for crystals bombarded with carbon ions without and with oxide films, respectively.

Two competing processes take place during deformation of single-crystal bismuth: twinning and slip.

Stimulating one of these processes leads to suppression of the other. This can be explained in terms of the conservation of energy. Let us denote the energy going into deformation of a solid by W_d and the energy of slip and twinning by W_{dis} and W_{tv}, respectively. Then, we can write

$$W_d = W_{dis} + W_{tv}. \qquad (2)$$

It is clear from this equation that a reduction in W_{dis} leads to an increase in W_{tv} and, conversely, an increase in W_{dis} leads to a reduction in W_{tv}.

One factor that suppresses slip may be impurities uniformly distributed in the crystal matrix.[6] Thus, twinning of pure iron by a static load can take place only at reduced temperatures.[7,8] Introducing several percent of silicon into the crystal makes slip so much more difficult that the samples deform mainly by twinning at room temperature under static loading.[9]

In the present experiments the factor inhibiting the development of slip was the implanted impurity carbon, which was localized in the surface layer of the irradiated material. An oxide film of thickness 100–1000 Å on the side facets of the crystal[1] leads to an increase in the stresses required to attain the same twinning rates as in irradiated samples without an oxide film. Since an oxide film cannot have much effect on the motion of twinning dislocations inside a crystal, it appears that the effect of a film is to slow down twinning dislocations near the surface.

CONCLUSION

A study of twinning in single-crystal bismuth irradiated by 25 keV carbon ions to a dose of 10^{17} ion/cm^2 and subjected to load pulses of duration $10^{-5} - 10^{-4}$ s with a stress amplitude of $(0.2-2.0) \times 10^3$ g/mm^2 has shown that ion processing suppresses slip and the activation of twinning in the crystals. Surface oxidation inhibits the movement of twinning dislocations.

[1] V. I. Bashmakov and M. M. Brodksiĭ, Fiz. Met. Metalloved. **35**, 163 (1973).
[2] V. V. Uglov, A. K. Kuleshov, B. Raushenbakh, A. Keniger, and K. Hammerl, *Materials from the IV All-Russia Conference on Modification of the Properties of Construction Materials by Charged Particle Beams* [in Russian], Tomsk (1996), p. 316.
[3] L. C. Lovell and J. H. Wernick, J. Appl. Phys. **30**, 234 (1959).
[4] V. I. Startsev, V. P. Soldatov, and M. M. Brodsky, Phys. Solid State **18**, 863 (1966).
[5] M. Hansen and K. Anderko, *Constitution of Binary Alloys*, McGraw-Hill, New York (1958).
[6] V. I. Bashmakov, N. G. Yakovenko, and L. F. Zavornaya, Fiz. Met. Metalloved. **29**, 947 (1970).
[7] I. A. Gindin and Ya. D. Starodubov, Fiz. Tverd. Tela **1**, 1794 (1959) [Sov. Phys. Solid State **1**, 1642 (1959)].
[8] R. I. Garber and I. A. Gindin, Usp. Fiz. Nauk **70**, 57 (1960) [Sov. Phys. Usp. **3**, 41 (1960)].
[9] D. Hull, Acta Metall. **8**, 11 (1960).

Translated by D. H. McNeill

Short-wavelength radiation by nonrelativistic particles

V. A. Buts

National Scientific Center, Kharkov Physicotechnical Institute, 310108 Kharkov, Ukraine
(Submitted March 30, 1998)
Zh. Tekh. Fiz. **69**, 132–134 (May 1999)

The possibility of exciting short-wavelength radiation from a flux of charged nonrelativistic particles is demonstrated analytically.   © *1999 American Institute of Physics.*
[S1063-7842(99)02505-2]

Today, ideas regarding the possibility of exciting short-wavelength radiation (up to x rays) from fluxes of charged particles mainly involve fluxes of relativistic particles. In this paper we show that nonrelativistic particles can efficiently generate short-wavelength radiation. This possibility occurs when charged particles interact with a periodic inhomogeneous medium, especially with crystals. The mechanism considered below is important because it is much easier to create nonrelativistic particle fluxes and oscillators than relativistic ones. Furthermore, their density can be much higher and is limited only by the density of solids. The last point is known to be decisive in creating the conditions for collective induced radiation.

We shall consider the emission from an oscillator whose trajectory can be written in the form $\mathbf{r} = \boldsymbol{\nu}_0 + \mathbf{r}_0 \sin \Omega t$ in a medium with dielectric constant $\varepsilon = \varepsilon_0 + q \cos \mathbf{k} \cdot \mathbf{r}$.

We shall assume that the velocities $\boldsymbol{\nu}_0$ and $\mathbf{r}_0 \Omega$ are not bounded and do not vary. There are many papers which deal with the emission from charged particles and oscillators in media with a periodic inhomogeneity. Mostly they consider radiation from relativistic particles, so that the analysis is muchsimpler. Of the papers which come closest to the case of interest to us below ($\lambda \gg d$, where d is the period of the inhomogeneity) we note the first paper on parametric Cerenkov radiation[1] and some papers on transition scattering.[2,3] In a study of the case $\lambda \gg d$ in Ref. 1, an average was taken: a layered inhomogeneous medium was replaced by an effective, in terms of the electrodynamics, anisotropic dielectric. It is easy to show that the radiation mechanism of interestto us below will vanish when this sort of average is taken. In papers on transition scattering, primary attention has been devoted to the case of greatest interest, scattering on a motionless charge ($M \rightarrow \infty$), and to the role of transition radiation in plasma physics. Besides, these papers did not discuss the radiation from oscillators. These papers contain all the elements required for solving the problem of interest to us. Using the method described in Ref. 2, together with the properties of Bessel functions, some simple, but cumbersome, transformations yield the following expression for the power radiated by an oscillator:

$$\frac{\partial W}{\partial t} = \left(\frac{eq}{c\varepsilon_0} \right)^2 \frac{1}{4\pi} \left\{ \sum_{n=1}^{\infty} \int d^3 k \right.$$

$$\times \left[|\omega_1| \delta\left(k^2 - \frac{\omega_1^2}{c^2} \varepsilon_0 \right) \frac{J_n^2(\mathbf{k}_{-1} \cdot \mathbf{r}_0) \mathbf{N}_{k-1,\omega}^n \cdot \mathbf{M}_{k-1}^n}{\left(k_{-1}^2 - \frac{\omega_1^2}{c^2} \varepsilon_0 \right)} \right.$$

$$\left. + |\omega_2| \delta\left(k^2 - \frac{\omega_2^2}{c^2} \varepsilon_0 \right) \frac{J_n^2(\mathbf{k}_{+1} \cdot \mathbf{r}_0) \mathbf{N}_{k+1,\omega}^n \cdot \mathbf{M}_{k+1}^n}{\left(k_{+1}^2 - \frac{\omega_2^2}{c^2} \varepsilon_0 \right)} \right]$$

$$+ \int d^3 k |\mathbf{k}_{-1} \cdot \boldsymbol{\nu}_0| \delta\left(k^2 - \frac{(\mathbf{k}_{-1} \cdot \boldsymbol{\nu}_0)^2}{c^2} \varepsilon_0 \right)$$

$$\times \left. \frac{J_0^2(\mathbf{k}_{-1} \cdot \mathbf{r}_0) \mathbf{N}_{k-1,\omega}^0 \cdot \mathbf{M}_{k-1,\omega}^0}{\left(k_{-1}^2 - \frac{\omega_1^2}{c^2} \varepsilon_0 \right)} \right\}, \tag{1}$$

where

$$\mathbf{M}_{\mathbf{k},\omega}^n = (\omega^2 \varepsilon_0 / c^2) \cdot \mathbf{L}_{\mathbf{k},\omega}^n - \mathbf{k}(\mathbf{k} \cdot \mathbf{L}_{\mathbf{k},\omega}^n);$$

$$\mathbf{N}_{\mathbf{k}\pm 1,\omega}^n = \mathbf{M}_{\mathbf{k}\pm 1,\omega}^n - \frac{\mathbf{k}}{k^2}(\mathbf{k} \cdot \mathbf{M}_{\mathbf{k}\pm 1,\omega}^n); \quad \mathbf{k}_{\pm 1} = \mathbf{k} \pm \boldsymbol{\kappa};$$

$$\mathbf{L}_{\mathbf{k},\omega}^n = \boldsymbol{\nu}_0 + n\Omega \mathbf{r}_0 / \mathbf{k} \cdot \mathbf{r}_0.$$

From Eq. (1), in particular, it is easy to obtain some of the standard results. In the following we shall examine two of the simplest cases where the features of the radiation mentioned above show up. Let an oscillator be at rest, $\boldsymbol{\nu}_0 = 0$, and $\boldsymbol{\kappa} \| \mathbf{r}_0 \| z$. Under these conditions, Eq. (1) yields

$$\frac{\partial W}{\partial t} = \left(\frac{e^2 \Omega^2 \cdot \beta_\perp^2}{3c} \right) \frac{3q^2}{2} \sum_{n=1}^{\infty} \frac{n^4}{m^2} J_n^2(m) \int_0^\pi (\sin \theta)^3 d\theta, \tag{2}$$

where $\beta_\perp = r_0 \Omega / c$ is the ratio of the velocity of the oscillator to that of light and $m = \boldsymbol{\kappa} \cdot r_0$ with $\kappa \gg k$.
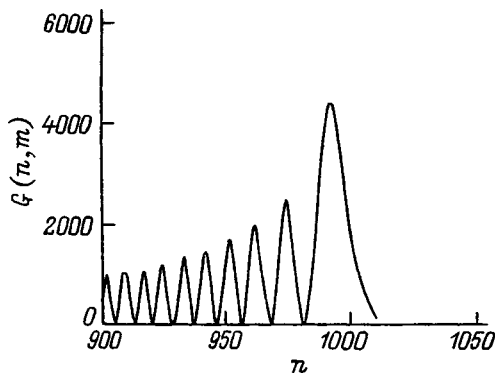
FIG. 1.

In Eq. (2) the first factor (the factor in parentheses) equals the dipole radiation power of a nonrelativistic oscillator in vacuum and each term of the sum describes the radiation at the corresponding harmonic n. The dependence of the radiated power on the harmonic number (n) and the degree of inhomogeneity (q) is given by the factor $G(m,n) = (q \cdot (n^2/m) J_n(m))^2$.

Using Eq. (2), let us compare the emission from an oscillator in vacuum and in a medium.[4] For this purpose, Fig. 1 shows the dependence of $G(m,n)$ on the harmonic number for $m = 10^3$ and $q = 10^{-3}$. This figure shows that the radiated power increases with n and has a maximum for $n = m$. Note that, in order to obtain a significant radiated power from an oscillator at harmonic number $n = 10^3$, this oscillator must have an energy $\gamma > 22$ ($\beta > 0.999$), while for a medium with a weak periodic inhomogeneity ($q = 10^{-4}$, $\varepsilon_0 = 1$), the same power can be produced by an oscillator with energy $\gamma = 1.0005$ ($\beta = 0.1$). As an illustration of the fact that in vacuum there is essentially no radiation from a nonrelativistic oscillator, while if a medium is present (even a very weak one), the radiation can be quite substantial, Fig. 2 shows a plot of the ratio of the power radiated by an oscillator in a medium ($q = 10^{-6}$) to the radiated power in vacuum for $\beta = 0.1$ ($m = 100$). It is clear that even a slight amount of periodically inhomogeneous medium ($q = 10^{-6}$) can make a fundamental change in the emission spectrum. Furthermore, the intensity of the radiation in the harmonics ($\omega = n\Omega$) can exceed the dipole radiation of the oscillator ($\omega = \Omega$). In fact, with Eq. (2) it is easy to find that, as long as the condition $2 \cdot [q \cdot (n^2/m) \cdot J_n(m)]^2 > 1$ is satisfied, the radiated power at harmonic n exceeds the dipole radiation from the oscillator
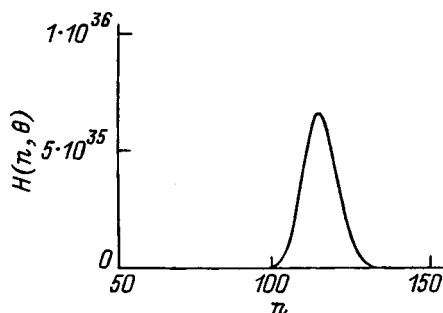
in vacuum. Given that the maximum value of the Bessel function is reached for large n at $n = m$, and using the asymptotic behavior of the Bessel function for $n \gg 1$ ($J_n(n) \propto n^{-1/3}$), this inequality simplifies to $(q \cdot n^{2/3}) > 1/\sqrt{2}$.

A high efficiency of ''long wavelength'' emission ($\lambda = d/\beta$) is characteristic of a nonrelativistic oscillator, but also of a nonrelativistic charged particle moving at constant velocity (ν_0) without oscillating ($r_0 = 0$) through a periodically inhomogeneous medium. For such a particle, the conservation laws imply that $\omega \approx \boldsymbol{\kappa} \cdot \boldsymbol{\nu}$. As in the case of the emission from an oscillator, we shall assume that $\boldsymbol{\kappa} \| \boldsymbol{\nu}_0 \| z$. In this case, the general formula (1) yields the following expression for the power radiated by the particle:

$$\frac{\partial W}{\partial t} = \left(\frac{e^2 \omega^2 \cdot \beta^2}{3c} \right) \frac{3q^2}{4\varepsilon_0^{3/2}} \int_0^\pi (\sin \theta)^3 d\theta, \qquad (3)$$

where $\beta = \nu_0/c$.

It is clear from this formula that the directional pattern of the radiation from a particle coincides with that from a motionless nonrelativistic oscillator. Furthermore, on comparing Eq. (3) with the formula for the emission from an oscillator in vacuum,[4] it is easy to see that if the quantity $\nu_0 \cdot q$ is substituted for the oscillator velocity in the latter formula, then it transforms to Eq. (3). Thus, a particle moving uniformly at velocity ν in a periodically inhomogeneous medium radiates as a motionless oscillator in vacuum in the dipole approximation, which oscillates at frequency ω while its oscillatory velocity is $\nu_{\mathrm{osc}} = \nu_0 \cdot q$.

We now estimate the prospects for this radiation in terms of the possibility of exciting radiation with a higher frequency. The minimum period of the inhomogeneity in a medium which can be used is the distance between atoms in a solid. It is of the order of $d = 10^{-8}$ cm. Noting the inequality $\lambda \gg d$, we can expect electromagnetic radiation with a minimum wavelength of $\lambda \approx 10^{-7}$ cm to be excited. In order to attain this purpose (radiation at a given wavelength), we must keep the equality $\omega = \boldsymbol{\kappa} \cdot \boldsymbol{\nu}$ in mind. Using this equality, we can find the magnitude of the oscillatory velocity which the oscillator must have, $\beta = \nu/c = \lambda/d$, i.e., in our case $\beta \approx 0.1$. If a charged particle becomes an oscillator as a result of an interaction with an external electromagnetic field with frequency Ω, then in order to attain an oscillatory velocity $\beta = 0.1$, the nonlinearity parameter of the wave, $\varepsilon = eE/mc\omega$ (where E is the electric field strength in the external wave), must equal 0.1.

Therefore, in the ideal limit we can calculate that when an electromagnetic wave with a wavelength of $\lambda = 10^{-4}$ cm and a nonlinearity parameter of 0.1 acts on a solid (crystal), the solid should emit radiation with $\lambda \approx 10^{-7}$ cm. We can also calculate that an analogous solution will be found when a beam of charged particles with velocity $\beta = 0.1$ is incident on a solid (crystal).

Here we have concentrated on the most interesting case of the excitation of x rays by an oscillator in a periodically inhomogeneous medium. It is clear, however, that this mechanism can show up in a considerably wider domain. For example, it can produce a high frequency component of the

FIG. 2.

emission spectrum from a plasma when periodic inhomogeneities are present within it or near its boundaries.

[1] Ya. B. Faĭnberg and N. A. Kizhniak, Zh. Éksp. Teor. Fiz. **32**, 883 (1957) [Sov. Phys. JETP **5**, 720 (1957)].

[2] V. L. Ginzburg and V. N. Tsytovich, Usp. Fiz. Nauk **126**, 553 (1978); V. L. Ginzburg, Usp. Fiz. Nauk **166**, 1033 (1996).

[3] V. L. Ginzburg and V. N. Tsytovich, *Transition Radiation and Transition Scattering* [in Russian], Nauka, Moscow (1983).

[4] A. A. Sokolov and I. M. Ternov, *The Relativistic Electron* [in Russian], Nauka, Moscow (1974).

Translated by D. H. McNeill

Skin effect during propagation of high-power short-wavelength microwaves in a partially ionized semiconductor plasma

O. A. Kosygin, V. N. Chupis, and A. Yu. Somov

Scientific-Research Institute of Mechanics and Physics at the N. G. Chernyshevski Saratov State University,
410071 Saratov, Russia
(Submitted May 15, 1998)
Zh. Tekh. Fiz. **69**, 135–136 (May 1999)

A study is made of the propagation of high-power electromagnetic waves in spatially inhomogeneous plasmas of thin semiconducting elements and films. The effect of surface recombination and an external magnetic field on the depth of penetration of an ionizing field into the semiconductor plasma is examined. © *1999 American Institute of Physics.*
[S1063-7842(99)02605-7]

Studies of the effect of the interaction of electromagnetic fields with ionized semiconductor plasmas contribute to the development of microwave technology at short wavelengths, in particular the creation of measurement and control devices for high power levels. Studies[1,2] have shown that stationary collisional ionization in semiconductors can be used for efficient control of the parameters of short-wavelength microwave radiation.

In this area, the greatest interest is in studies of the interaction of strong electromagnetic waves with spatially inhomogeneous plasmas in thin (with thicknesses less than the characteristic diffusion length) semiconducting elements.[3,4] In these semiconductors the diffusion of highly energetic carriers from the heating region has a significant effect on the magnitude of the threshold fields and electrodynamic characteristics of a thin semiconducting element.[5] For this reason, in thin semiconductors a number of new phenomena have been observed that are associated with the effect of the surface on the carrier distribution within the bulk semiconductor and, accordingly, on the interaction of strong electromagnetic fields with thin semiconductors.[6] A theoretical and experimental analysis has been made[3] of the skin effect during the propagation of short-wavelength ionizing radiation in semiconductor plasmas. In particular, it was found that the penetration depth of an ionizing field in semiconductors is essentially independent of its amplitude at the surface, but is determined by the physical characteristics of the semiconducting material. There is considerable interest in studies of the dependence of the penetration depth of strong fields in semiconductors on the properties of the surface, in particular on the surface recombination rate. By analogy with the earlier work,[4,6] one might expect that surface recombination has a significant effect on the way high-power short-wavelength microwaves propagate.

In this regard, in this paper we study the dependence of the penetration depth of an ionizing field into a semiconductor on the surface recombination rate and the strength of an applied magnetic field.

The system of equations describing the interaction of an ionizing field with a semiconductor in this case has the form

$$\frac{d^2 E}{dz^2} - \frac{\omega_p^2}{c^2}\frac{n}{n_0} E = 0, \quad D\frac{d^2 n}{dz^2} + \nu_i n - \nu_r n - \nu_r n_0 = 0,$$

$$\frac{\nu_i}{\nu_r} = G \exp\left\{-\left(\frac{E^*}{E}\right)^2\right\}, \tag{1}$$

where E is the strength of the electromagnetic field in the semiconductor, ω_p is the plasma frequency, c is the speed of light, D is the diffusion coefficient, and ν_i and ν_r are the average ionization and recombination rates.

G and E^*, which depend on the properties of a specific semiconductor and the frequency of the ionizing microwave field, can be determined experimentally or from kinetic theory. For example, for InSb at $f = 37$ GHz, $G = 9.8 \times 10^4$ and $E^* = 1.2 \times 10^3$ V/cm.[6] In solving the coupled system of Eqs. (1), the initial values of the electromagnetic field-strength and carrier concentration at the boundary of a semiconducting element were specified. The semiconducting slab was broken up into N parallel layers in the z plane and the concentration was assumed constant within each layer, while the tangential components of the electromagnetic field strength were matched at the boundaries of the layers. The calculated $E(z)$ curves for InSb ($n_0 = 2.8 \times 10^{12}$ cm^{-3}, $T = 77$ K) are shown in Fig. 1.

The calculations show that surface recombination changes the way the electromagnetic wave propagates. In this case, because of charge carrier ''depletion'' in the surface layer, the ionizing field penetrates considerably more deeply into the semiconductor plasma (curve *1* of Fig. 1).

Figure 2 shows plots of the penetration depth δ_{HF} of an electromagnetic field into a semiconductor plasma for different rates P_s of surface recombination.

In this case, as the surface recombination rate increases, the concentration of nonequilibrium carriers near the semiconductor surface decreases, along with the plasma frequency, and this leads to a rise in the depth of penetration of the ionizing field into the plasma.

As in Ref. 4, here an external magnetic field in a Voigt geometry, $B \uparrow \uparrow E \perp k$, makes it possible to reduce the influence of the surface on the propagation of an ionizing wave in a semiconductor plasma (Fig. 2). The physical nature of this
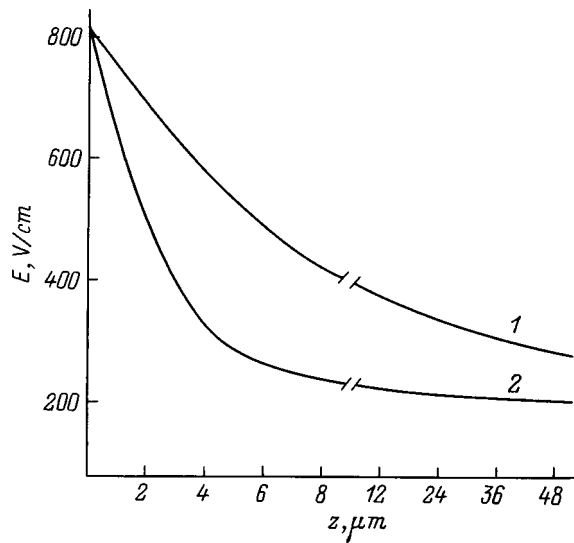
FIG. 1. Field dependences of the depth of penetration of ionizing radiation into a semiconductor plasma: $P_s = 10^6$ cm/s (*1*), 0(*2*).

FIG. 2. Plots of the penetration depth of ionizing radiation into the plasma: $B = 0$ (*1*), 0.1 (*2*), 0.2 (*3*), 0.3 T (*4*).

phenomenon is that, with a higher magnetic field the electrons are ''twisted'' around the lines of force and the ionized plasma is ''squeezed away'' from the semiconductor surface, which is equivalent to a reduction of the characteristic diffusion length in an external magnetic field. Because of this, for a magnetic induction $B > 0.2$ T, surface recombination essentially has no effect on the depth of penetration of an ionizing field into a semiconductor (curves *2–4* of Fig. 2).

These features of the interaction of high-power electromagnetic radiation with thin semiconductors, especially the dependence of ω_p and the penetration depth of the electromagnetic field into a semiconductor plasma on the external magnetic field, are of fundamental importance for thin semiconductors and s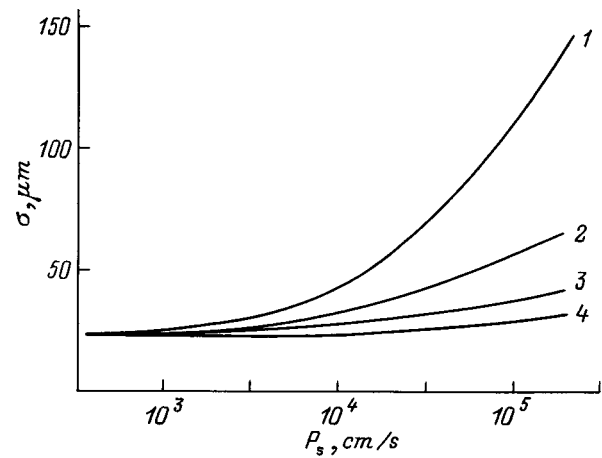hould be taken note of during the develop-ment of control and measurement devices for short-wavelength microwaves at high power levels.

[1] F. G. Bass and Yu. G. Gurevich, *Hot Electrons and High-Power Electromagnetic Waves in Semiconductor and Gas Discharge Plasmas* [in Russian], Nauka, Moscow (1975), 368 pp.

[2] D. A. Usanov, B. B. Feklistov, and A. Yu. Vagarin, Radiotekh. Elektron. **24**, 1681 (1979).

[3] V. N. Chupis, O. A. Kosygin, N. A. Dukhovnikov, and E. M. Semenova, Pis'ma Zh. Tekh. Fiz. **19**(23), 69 (1993) [Tech. Phys. Lett. **19**(12), 764 (1993)].

[4] V. N. Chupis, A. Yu. Somov, O. A. Kosygin, and E. M. Semenova, Pis'ma Zh. Tekh. Fiz. **21**(10), 16 (1995) [Tech. Phys. Lett. **21**(5), 358 (1995)].

[5] V. N. Chupis and L. I. Kats, Fiz. Tekh. Poluprovodn. **17**, 1288 (1983) [Sov. Phys. Semicond. **17**, 814 (1983)].

[6] E. Yu. Al'tshuller, L. I. Kats, and V. N. Chupis, Radiotekh. Elektron. **37**, 560 (1992).

Translated by D. H. McNeill

Determination of the temperature of the onset of ion transport from the temperature dependence of the relative change in the ultrasound velocity in the lithium ceramic Li–Si–Ge–As–S–O

Yu. F. Gorin, O. L. Kobeleva, V. L. Kobelev, N. V. Mel'nikova, Ya. L. Kobelev, L. Ya. Kobelev, and O. S. Tsyganov

A. M. Gorki Urals State University, 620083 Ekaterinburg, Russia

An ultrasonic method is used to study the onset of ion transport inlithium oxides in a ceramic of the Li–Si–Ge–As–S–O system. The onset temperature for ion transport estimated from the temperature dependence of the relative change in the propagation speed of ultrasound is the same as that obtained by other methods, which indicates that the ultrasonic method is applicable to studies of the onset of ion conductivity inlithium oxide semiconductors. © *1999 American Institute of Physics.* [S1063-7842(99)02705-1]

An important task in research on new ion semiconductors is to establish the temperature at which ion transport becomes significant. The methods used for estimating the onset temperature of ion transport in research on the complex chalcogenides of silver and copper from the temperature dependences of the electrical conductivity and dielectric permittivity,[1,2] from an examination of the electrical conductivity in cells with filters, and from nuclear magnetic resonance (NMR) determinations of the ion mobility can be supplemented by other techniques, such as studies of the temperature dependence of the relative change in the velocity of ultrasound in these compounds. The feasibility of determining the onset temperature for ion transport in silver and copper chalcogenides by studying the temperature dependence of the relative change in the ultrasound speed has been demonstrated.[3] It has not been clear whether the complex chalcogenides of silver and copper are the only mixed ionic semiconductors for which the ultrasound technique is useful for determining the ion transport onset temperature.

This paper is devoted to using the ultrasound technique for studying the onset of ion transport in lithium oxides. (We consider a new ceramic which we have synthesized in the Li–Si–Ge–As–S–O system as the test material.)

Impedance studies of ceramic Li–Si–Ge–As–S–O at temperatures of 220–520 K, along with studies of the dielectric permittivity and of the temperature dependence of the relaxation time of the magnetic moment of Li^7 nuclei, have shown that this compound is a mixed (electronic–ionic) semiconductor with lithium ion conductivity. The fraction of ion conductivity is 48% at 300 K and 70% at 450 K.[4]

The temperature dependences of the conductivity and dielectric permittivity were studied at temperatures of 220–650 K and a frequency of 1.592 kHz, which lies in the frequency region where the effect of electrode processes on the characteristics being studied can be neglected. (This region was determined by analyzing the frequency dependences of the impedance and admittance of this compound.) Figure 1 shows the specific electrical conductivity of the oxide semiconductor Li–Si–Ge–As–S–O as a function of temperature. An analysis of this dependence showed that the activation energy at 220–240 K is 1.03 eV and for 240–350 K it is 0.56 eV. The change in the activation energy at $T \approx 240$–250 K, together with data on the mobility of the Li^7 ions (the spin–lattice relaxation rate of Li^7 ions has been studied[4,5] at temperatures of 240–650 K), are indicative of a change in the mechanism for conductivity and of the onset of lithium ion conductivity. With the onset of significant ion conductivity, the dielectric permittivity begins to rise rapidly. The range of temperatures corresponding to the onset of the rapid rise in the dielectric permittivity (250–260 K)[4] is consistent with the range of temperatures where the activation energy changes.

These studies of the acoustic properties of the ceramic were carried out by the phase-pulse method using a quartz ultrasonic generator (the frequency of the longitudinal wave was 5 MHz and that of the transverse wave, 2.5 MHz) as a source of ultrasound. The temperature dependences of the relative changes in the propagation speeds of longitudinal and transverse ultrasound in a Li–Si–Ge–As–S–O sample
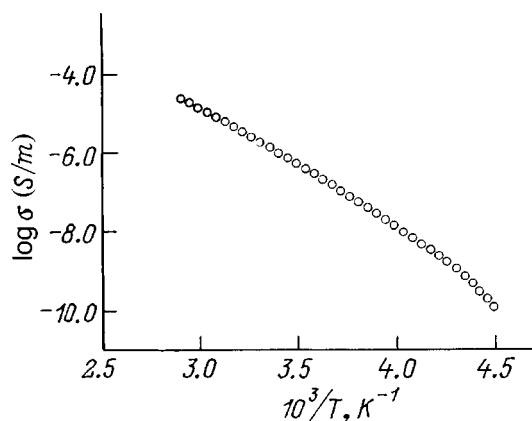

FIG. 1. Specific electrical conductivity of Li–Si–Ge–As–S–O oxide semiconductor as a function of temperature.
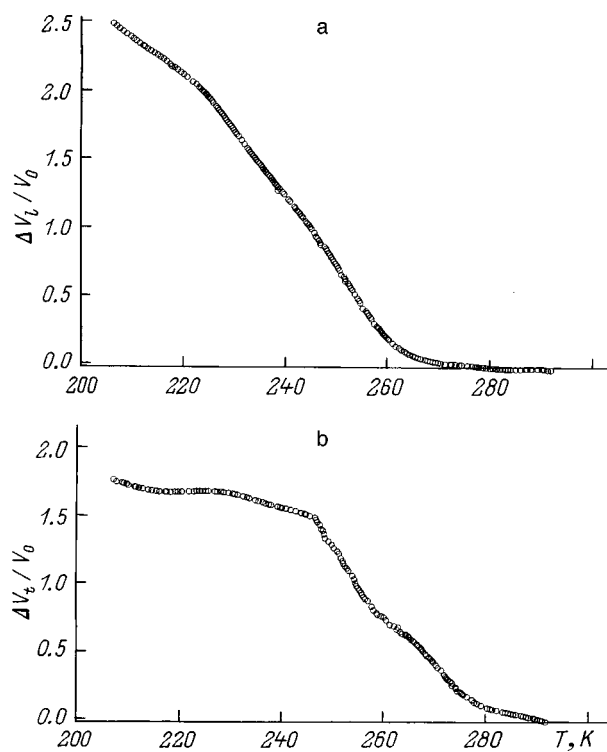
FIG. 2. The relative changes in the propagation velocities of longitudinal (a) and transverse (b) ultrasound in a sample of Li–Si–Ge–As–S–O as functions of temperature.

are shown in Fig. 2. These curves typically have three temperature intervals where significant changes in the relative ultrasound speed take place: 215–220, 240–250, and 260–280 K. The changes in the slope of the plots of the relative change in the ultrasound speed within these temperature intervals are evidence of structural changes in the crystal lattice at these temperatures which can be interpreted as phase changes. The more rapid fall in the relative change in the sound speed for $T > 250$ K is related to the appearance of free lithium ions. The temperature where the ultrasound speed changes sharply (Fig. 2b), 250 K, can be regarded as the ion transport onset temperature.

The ion transport onset temperature estimated from the temperature dependence of the relative variation in the ultrasound propagation speed is, therefore, in agreement with the value obtained by the impedance method (240–250 K)[4] and with an estimate of the ion transport onset temperature obtained from an NMR study of the mobility of the Li^7 ions.[4,5]

These results indicate that the ultrasonic method can be used to study the onset of ion conductivity in oxide ionic semiconductors.

[1] V. B. Zlokazov, N. V. Mel'nikova, E. R. Baranova *et al.*, Elektrokhimiya **28**, 1523 (1992).
[2] E. R. Baranova, V. B. Zlokazov, L. Ya. Kobelev, and M. V. Perfil'ev, Pis'ma Zh. Tekh. Fiz. **16**(10), 27 (1990) [Sov. Tech. Phys. Lett. **16**(5), 372 (1990)].
[3] Yu. F. Gorin, N. V. Mel'nikova, E. R. Baranova, and O. L. Kobeleva, Pis'ma Zh. Tekh. Fiz. **23**(14), 35 (1997) [Tech. Phys. Lett. **23**(7), 550 (1997)].
[4] Ya. L. Kobelev, V. B. Zlokazov, O. L. Kobeleva *et al.*, *Scientific Proceedings of the All-Russia Conference on the Physics of the Condensed State* [in Russian], Sterlitamak (1997), pp. 102–104.
[5] O. L. Kobeleva *et al.*, Joint International Meeting of ECS and ISE, Paris (1997), Poster Q1-2156.

Translated by D. H. McNeill

Signal amplitude and amplitude spectrum shape of a detector with incomplete charge transport

N. B. Strokan

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted May 12, 1998)
Zh. Tekh. Fiz. **69**, 139–142 (May 1999)

A case is considered in which, unlike the high-energy-resolution regime, the ratio of the size of
the working zone of the detector to the drift displacement length of the charge carriers
cannot be assumed small ($\ll 1$). This analysis pertains to detectors based on semi-insulating
GaAs (SI GaAs) in the detection of short-range ions. It is assumed that most carrier capture occurs
during the drift of the carriers. The expressions derived for the signal amplitude and the
shape of its spectrum are shown to contain three independent parameters. When these are
determined experimentally, it is possible to find the average field strength and its velocity
of propagation in the structure as functions of the applied voltage, and also to find the carrier
lifetime to capture and the degree of nonuniformity of the capture over the volume of
the working zone of the detector. © *1999 American Institute of Physics.*
[S1063-7842(99)02805-6]

In recent years problems have arisen in high energy physics which can only be solved with semiconductor detectors capable of operating in intense radiation fields. In this regard, detailed studies are being made of the radiation hardness of detectors based on high resistance silicon, the technology of which is well developed. In addition, the prospects for new materials, especially SI (semi-insulating) GaAs, are being examined. Low conductivity is attained in this material by compensation of donor and acceptor centers. Ultimately, both irradiated (initially pure) Si and initial SI GaAs are characterized by the presence of a substantial number of charge-carrier capture centers in the detector volume. This leads to incomplete transport of the nonequilibrium charge created by the nuclear radiation.

In this paper we examine the behavior of the main detector characteristics, the average signal amplitude and the amplitude spectrum, when capture centers appear in the bulk of the material. The calculation has been done for a detector based on SI GaAs and for the detection of short-range radiation.

1. The typical geometry of an experiment, where radiation enters from the side of the p^+ contact, is illustrated in Fig. 1. The detector structure is partially depleted, to the plane W with an overall extent d. The distribution of electron-hole pairs in a track (of length $R < (W,d)$) is specified by the function $G(y)$. It is assumed that of the two channels for charge loss, recombination directly in the particle track or localization during drift, the second mechanism predominates.[1] With localization, the signal is determined by Faraday's law, as well as by the type of capture. Ultimately, the charge transferred in a region with an electric field E is given by

$$q_0 = dx/WG(y)\exp(-dx/\mu E\tau)dy. \qquad (1)$$

This equation implies that electrons from the layer y

move at a velocity μE (where μ is the current carrier mobility) and are subject to capture with a characteristic time τ. The contribution to the signal in the plane W is determined[2,3] by the potential difference $\Delta V(x)$ traversed by an electron before capture, where $V(x)$ is the potential profile in the capacitor with a geometry (plane, cylindrical, spherical) equivalent to the detector for a volume charge of zero. In our case, this reduces to the fraction of the path traveled out of the interelectrode separation $[0,W]$.[1]

We shall proceed further with the model for formation of a field region in SI GaAs proposed in Ref. 5. It is characterized by an extremely weak dependence of the field E on position. This makes it possible to set the drift transport mean free path $\mu E\tau$ constant in the field region. We also take into account that $R < W$, d. Then the signal in the plane W, normalized to the charge implanted by the alpha particles, is

$$q = (\mu E\tau/W)[1 - \exp(-W/\mu E\tau)](\mu E\tau/R)$$
$$\times[\exp(R/\mu E\tau) - 1]. \qquad (2)$$

For an SI GaAs detector, this expression must be multiplied by a factor W/d, which accounts for the slow relaxation of the semi-insulating base of the structure. Finally, for $R \to 0$ we obtain the detector signal

$$q = (\mu E\tau/W)[1 - \exp(-W/\mu E\tau)](W/d). \qquad (3)$$

Since in an experiment one records the dependence of the signal on the voltage U at the detector, the function $W = f(U)$ determines the final form of Eq. (3). In this form of the model,[5] $W = \gamma U$, where $1/\gamma = E_{av}$ is the average magnitude of the field and Eq. (3) transforms to

$$q(U) = (\mu\tau/\gamma d)[1 - \exp(-\gamma^2 U/\mu\tau)]$$
$$= P_1[1 - \exp(-P_2 U)], \qquad (4)$$

where the parameters $P_1 = \mu\tau/\gamma d$ and $P_2 = \gamma^2/\mu\tau$.
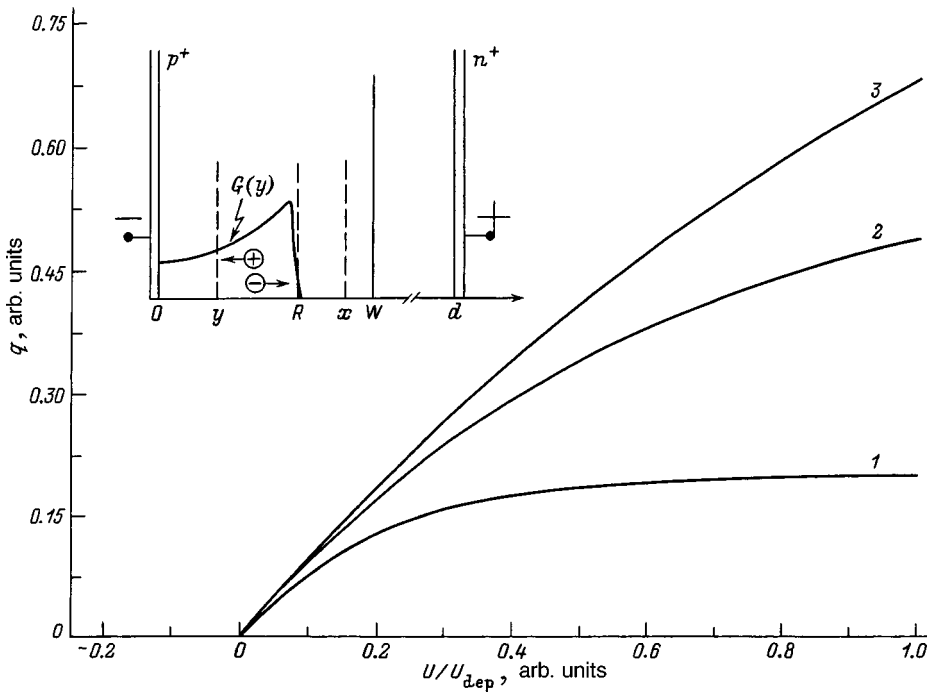
FIG. 1. Amplitude of the detector signal as a function of bias voltage: $d/\mu E\tau = 5.0$ (*1*), 1.66 (*2*), 0.83 (*3*); $\tau = 1.0$ (*1*), 3.0 (*2*), and 6.0 ns (*3*). The inset shows the geometry of the structure under consideration.

When U is increased to depletion of the structure, it is necessary to set $W = d$ in Eq. (3). In addition, depletion is observed in practice at substantial fields E, where the drift velocity saturates to $v_s = \mu E$. Under these conditions $q(U)$ vanishes and the signal reaches its limit

$$q_{\max} = (v_s\tau/d)[1 - \exp(-d/v_s\tau)]. \tag{5}$$

Figure 1 shows the charge as a function of detector voltage normalized to the depletion value $U_{\text{dep}} = d/\gamma$. In the calculations with Eq. (4) we have used typical practical parameters: $\gamma = 10^{-4}$ cm/V, $d = 5 \times 10^{-2}$ cm, and $v_s = 10^7$ cm/s. The parameter $d/\mu E\tau$ was set to 0.83, 1.66, and 5.0, corresponding to lifetimes of 6.0, 3.0, and 1.0 ns. It is clear that, as $d/\mu E\tau$ increases, the function $q(U)$ approaches saturation. At the same time, the maximum realizable amplitude is a smoothly decreasing function of $d/\mu E\tau$ (see Eq. (5) and the inset to Fig. 3). In practice it is important that, in approximating the experimental data with Eq. (4), we can find the quantities P_1 and P_2, as their product $P_1 \times P_2 = \gamma/d$ gives the value of $(U_{\text{dep}})^{-1}$ and of the average field strength $1/\gamma = E_{\text{av}}$. In addition, the values of τ are determined from P_1.

2. We now obtain an expression for the shape of the amplitude spectrum, $dN/dq = f(q)$, for this model. To do this it is necessary to explain the scatter in the values of q. It is logical to attribute the scatter to fluctuations over the volume of the detector in the time τ to carrier capture as a parameter which is more sensitive to the structural perfection of the material.[2] Then we pick a Gaussian distribution for τ. That is the most probable form when the deviation from the average value of a variable quantity is less than the mean itself, i.e., $|\tau - \tau_0| < \tau_0$, so that

$$dN/d\tau = \exp[-(\tau - \tau_0)^2/2\sigma^2]/\sqrt{2\pi}\sigma, \tag{6}$$

where σ is the dispersion.

We introduce the half width of the distribution of τ relative to the average value, $R_\tau = 2.35\sigma/\tau_0$. Then, omitting the constant terms, for the spectrum shape we obtain (P_2 corresponds to the τ_0 and $z = \tau/\tau_0$)

$$
\begin{aligned}
dN/dq &= (dN/d\tau)/(d\tau/dq) \\
&= \frac{\exp[-(z-1)^2/0.362R_\tau^2]}{1 - [1 + P_2(U/z)]\exp[-P_2(U/z)]}.
\end{aligned} \tag{7}
$$

When the structure is depleted and the charge obeys Eq. (5), the numerator of Eq. (7) is the same as before, but the denominator is given in terms of P_1 as $[1 - (1 + 1/zP_1)\exp(-1/zP_1)]$. The shape of the spectrum (like the average signal magnitude) no longer depends on the bias U.

Plotting the spectra showed that the shape of the spectral line according to Eq. (7) is sensitive to the degree of nonuniformity of the capture time R_τ. Thus, for $R_\tau \approx 1$ the line is asymmetric and has an extended left side. As R_τ is reduced, the line becomes more symmetric. The quantitative characteristic of the line shape was taken to the its half width. It turns out that as the bias on the detector is raised, the linewidth also increases and reaches saturation when the structure is depleted (Fig. 2). This last point is in conflict with the practice for spectrometric detectors with near complete charge transport. Thus, there is a nearly parabolic rise in the relative resolution as a function of the induced charge (inset to Fig. 2).

Figure 3 traces the dependence of the linewidth and amplitude on the parameter $d/\mu E\tau$ when the structure is depleted. As noted above, $q_{\max}$ typically drops significantly, while the linewidth varies little over the interval $d/\mu E\tau = 1 - 5$. The linewidth, however, depends significantly on the charge-carrier capture nonuniformity (R_τ). This makes it
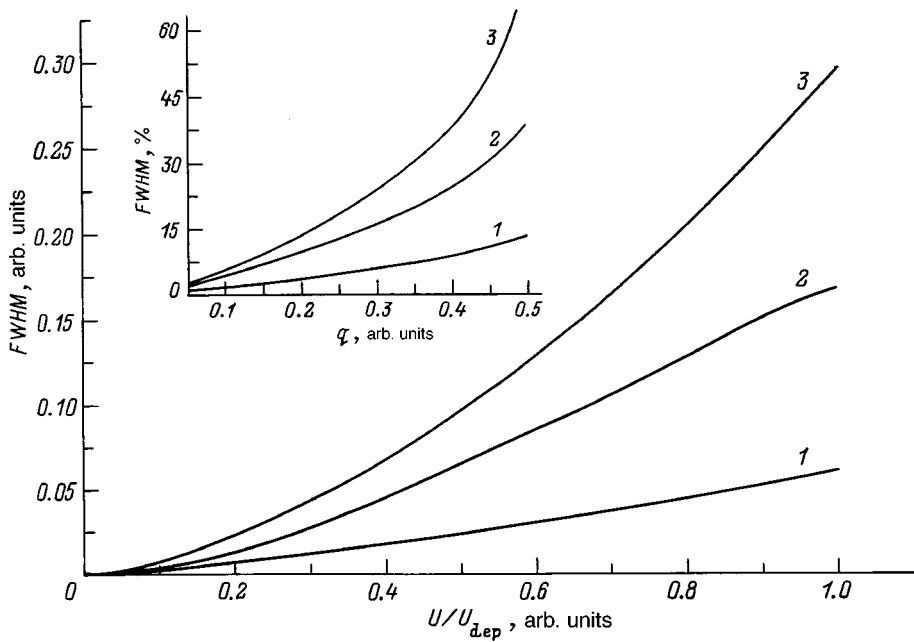
FIG. 2. The spectrum linewidth as a function of detector bias voltage: $R_\tau = 0.2$ (*1*), 0.6 (*2*), 1.2 (*3*); $\tau = 3$ ns, $d/\mu E\tau = 1.66$. The inset shows the linewidth relative to the amplitude as a function of the signal amplitude.

possible to determine R_τ for a known ratio $d/\mu E\tau$ by comparing the measured linewidth with the data of Fig. 3.

3. We conclude by noting what we believe to be the main results. We have examined the operation of a detector with rapid capture of charge carriers through localization. We have derived expressions for the spectrum line shape and average signal amplitude as functions of the detector bias.

These equations contain three independent parameters which are determined in the course of an experiment. By comparing the computations and experimental data, it is possible in principle to determine the structure of the electric field in the detector and the carrier transport characteristics of the material. These include the extent of the field region

and the average field strength, the lifetime to localization at capture centers of the drifting carriers, and the degree of nonuniformity of the capture in the detector volume.

We have pointed out differences from the case of weak capture. The main difference is a rise in the spectrum linewidth with detector bias voltage (as the signal amplitude increases).

If capture occurs, as does carrier recombination, in a ''track plasma'' state, then after complete depletion of the detector the signal should rise with the detector bias. Here the charge losses will decrease with the field strength E as $1/E$ for light ions and as $\ln(1/E)$ for heavy ions.[7]

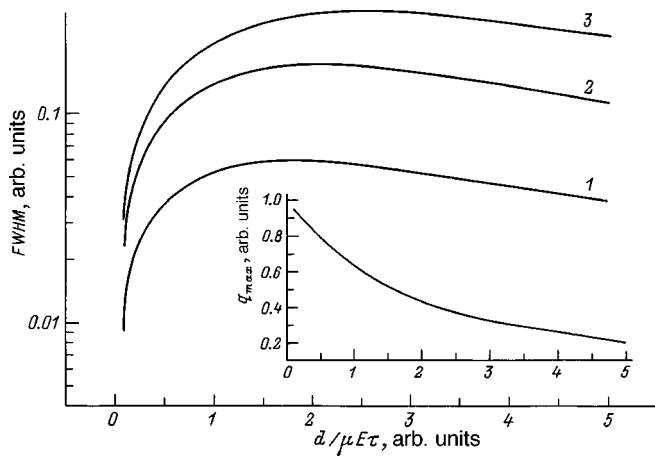The author thanks E. M. Verbitskaya, V. K. Eremin, and A. M. Ivanov for valuable comments.

FIG. 3. The spectrum linewidth as a function of the ratio of the sizeof the field region to the drift displacement length: (*1–3*) as in Fig. 2. The inset shows the same for the amplitude of the induced charge.

[1])The structure of Eq. (1) was first proposed in Ref. 4, but the authors made an error in their treatment of Faraday's law by considering a volume charge in the variation of $V(x)$.

[2])The question of nonuniformity of the capture was first discussed by Makovsky *et al.*[6]

[1]L. L. Makovskiĭ, N. B. Strokan, and N. I. Tisnek, Fiz. Tekh. Poluprovodn. **2**, 972 (1968) [Sov. Phys. Semicond. **2**, 807 (1968)].

[2]G. Gavalleri, E. Gatti, G. Fabri, and V. Svelto, Nucl. Instrum. Methods **92**, 137 (1971).

[3]V. K. Eremin, S. G. Danengirsh, N. B. Strokan, and N. I. Tisnek, Fiz. Tekh. Poluprovodn. **8**, 556 (1974) [Sov. Phys. Semicond. **8**, 355 (1974)].

[4]G. L. Miller and W. M. Gibson, Nuclear Electronics **1**, 477 (1962).

[5]D. S. McGregor, R. A. Rojeski, G. F. Knoll *et al.*, Nucl. Instrum. Methods Phys. Res. A **343**, 527 (1994).

[6]L. L. Makovsky, N. B. Strokan, and N. I. Tisnek, IEEE Trans. Nucl. Sci. **NS-15**, 304 (1968).

[7]V. Eremin, I. Ilyashenko, N. Strokan, and B. Schmidt, Nucl. Instrum. Methods Phys. Res. A **377**, 184 (1996).

Translated by D. H. McNeill